TENSORIAL MATRICES

Jerzy Plebanski*
Departamento de Física, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México*

(Recibido: 17 julio 1964)

ABSTRACT

The methods of linear algebra are applied to the algebraic analysis of the structure of objects possessing arbitrary sets of indices. With the help of the methodics of analytic functions, general relationships are established between invariants of various types (similar to Rodríguez' formula). The theory of meromorphic functions on generalized matrices is developed, and the theory of canonical forms of generalized matrices is outlined. The theory developed can be used in the study of the algebraic structure and of the algebraic types of the arbitrary spinor-tensor fields. (E.g. spinorial fields with arbitrary number of indices, lorentzian and isotopical, various curvature objects, etc.). Thus, possible applications can be found in field theory (classical and quantum), and in particular in general relativity.

*On leave of absence from the University of Warsaw, Warsaw, Poland.

I. INTRODUCTION

In the various fields of theoretical physics one often meets quantities (in general complex) which are simultaneously tensors with respect to a few groups, e.g., C_∞, Lorentz group, isotopical group, etc. As such, these quantities (fields) 'wear' *sets* of indices referring to the transformations related to the given irreducible representations of the groups in question. The *sub-sets* of the indices of the definite type referring to the given representation of the given group may be submitted to some symmetry requirements.

The recent developments in the methods of theoretical physics are often related to the study of the algebraic properties of such quantities. (E.g., in general relativity the Petrov–Penrose classification of the algebraic structure of the conformal curvature has stimulated important progress in the covariant study of the asymptotics, in the theory of the gravitational radiation as well as in the theory of exact solutions; see e.g., [1], [2], [3], [4], [5]).

Usually one studies the algebraic properties of the objects mentioned above by applying rather standard techniques of the linear algebra: the canonical Jordan's classification of matrices, studies of various types of the eigen-values problems, and specific spinorial techniques ([6], [7], [8]).

Nevertheless, in each specific problem one usually has to construct 'ad hoc' a special method of applying these techniques of the linear algebra – a method which would be manifestly covariant with respect to the groups to which the indices of the studied field refer. In such adaptations of the methodics of the linear algebra the tensorial notation with all indices specified explicitly is somewhat confusing. Sometimes therefore, one applies a more compact notation, e.g., the bi-vectors formalism [9]. However, such an improved notation usually has to be adopted "ad hoc" in specific problems; there seems not to exist a generally accepted method of introducing it.

The aim of this study is to approach the mentioned problems in an unified way, to explore the general situation and to introduce a convenient adaptation of the methods of the linear algebra fitted to it. A simple notation is proposed which forms a sort of compromise between the tensorial and the matrix notations. It shares the advantages of both: it is compact similarly to the matrix notation; similarly to the tensorial notation it is manifestly covariant and allows the convenient use of the techniques operating with the Levi-Civita's and Kronecker's symbols.

It is to be stressed that we do not intend to present here any essentially new mathematical results. In fact, all algebraic problems which one meets in theoretical physics usually have solutions implicitly hidden in textbooks of linear algebra. What we intend to do in this paper consists simply in the application of the well-known ideas and methods to the specific general situation which one meets in the problems of theoretical physics. This situation, however, we intend to explore in a systematical way.

For the reasons of compactness and self-consistency we shall sketch ideas of proofs of the mathematical facts we use. Most of these facts, however, follow from more profound general theorems of the linear algebra. The reasons why we care to sketch our more primitive proofs are simply: (1) we want to illustrate our techniques and notation especially those based on applications of analytic functions (2) this paper is thought as providing the theoretical physicist with some working tools which he may use without looking into mathematical references; it is better to convince him that the quoted theorems and results are true.

II. SETS OF INDICES, NOTATION

Let the theory we deal with be covariant with respect to some set of groups $\mathcal{G}_1$, $\mathcal{G}_2$, ... Suppose that we characterize these groups by their irreducible representations. Therefore, the tensorial quantities appearing in such a theory shall 'wear' sets of indices referring to the various representations of the groups

Let various possible kinds of indices be denoted as:

$A' \to 1', 2', \ldots, n'$, $A'' \to 1'', 2'', \ldots n''$, etc. [capital Roman types; primes, bises, etc., distinguish different kinds of indices. It may happen, however, that A', A'' refer to different representations of the same group while A''' refers to some other group]. The capital Roman indices may appear as co-or contra-indices which, respectively, are assumed to transform co-or contra-gradiently under the transformations of the given representation of the given group. Therefore the contraction (the Einstein's summational convention will be assumed) of the indices of the same kind is an invariant operation. The groups of indices of the same kind may possess some definite symmetries-which is also a tensorial property.

We would like to introduce an abbreviated notation: a, b, c. . . (lower-case Roman types) will stand for *sets* of capital Roman indices with definite symmetries (in particular, no symmetries), assumed for any sub-set of capital Roman indices of the same kind (or, more generally, for any sub-set of A', A'', . . ., all referring to a *single* given group). For the definite type of the symmetrization - separate for each kind of indices - we shall use the general symbol $\{ \quad \}$. Therefore, we identify

$$a \equiv \{A'_1, \ldots, A'_p \; ; \; A''_1, \ldots, A''_q \; ; \ldots \}. \tag{2.1}$$

The sets of indices a, b, . . . may appear as well as co-or contra-sets attached to complex fields*.

Our set of indices a runs through its *values* $\cancel{a}$, which consist of all possible substitutions of numbers in place of the capital Roman indices. These values we group into the ensembles of the *essentially different values* $\overset{1}{\cancel{a}}, \overset{2}{\cancel{a}}, \overset{N}{\cancel{a}}$.

By this we mean, that if an object T_a has the independent components $\overset{1}{T}, \overset{2}{T}, \ldots, \overset{n}{T}$ then

*All further theory applies when the studied objects have components from any algebraically closed field of numbers; for simplicity, in this text we restrict ourselves to complex objects.

for $\underset{\sim}{a}$ from $\overset{k}{\not a}$ $$T_a = \pm \overset{k}{T}, \quad k = 1,2,\ldots,N. \tag{2.2}$$

[Plus or minus to leave place for the skew symmetries]. The number of numerical values which constitute $\overset{k}{\not a}$ we denote by p_k and we will call the weight of the essentially different value $\overset{k}{\not a}$. It is obvious that the number N in (2.2) which is characteristic for our set a may be understood as the highest of the numbers k for which are possible non-trivial objects of the type $T_{a_1 \ldots a_k} = T_{[a_1 \ldots a_k]}$. The square bracket denotes the total symmetry with respect to the sets $a_1, \ldots, a_k$. (It does not affect the internal skew symmetries) of sets $a_1, \ldots, a_k$ in their capital Roman indices).

We will illustrate these definitions by two simple examples:

(1) Let $a \equiv [\alpha_1 \alpha_2]$, $\alpha = 0,1,2,3$ refers to $C\infty$; [ ] stands for the skew symmetry. Obviously $\overset{1}{\not a} = ([0,1], [1,0])$, $\overset{2}{\not a} = ([0,2], [2,0]), \ldots$ $\overset{6}{\not a} = ([1,2], [2,1])$. Here the N is equal to 6 and all the p_k are equal to 2.

(2) Let $a \equiv (A_1 A_2 \ldots A_{2s})$, $A = 1,2$ is a spinorial index (referring to the unimodular group), round bracket stands for total symmetrization, $s = 1/2, 1, 3/2,$ Here we have

$$\overset{1}{\not a} = (11\ldots), \ldots, \overset{k}{\not a} = (\underbrace{1\ldots1}_{2s-k+1} \quad \underbrace{2\ldots2}_{k-1}), \text{ in } p_k = \binom{2s}{k-1} \tag{2.3}$$

combinations, $\ldots \overset{2s+1}{a} = (22\ldots)$.

Here the number N is $2s+1$.

These examples are rather trivial, but they show clearly the construction of our definitions. The aim of this paper is precisely to provide ourselves with the mathematical machinery capable of handling conveniently much more complicated sets of indices, by the use of the concepts so naively simple in examples mentioned.

III. TENSORIAL MATRICES, MULTIPLICATION

By the *square* N × N tensorial matrix we understand the collection of complex numbers:

$$M^{a}_{b} \underset{df}{=} M^{\{A'_1 \dots A'_p \,;\, A''_1 \dots A''_q \,; \dots\}}{}_{\{B'_1 \dots B'_p \,;\, B''_1 \dots B''_q \,; \dots\}}. \tag{3.1}$$

where the co- and contra-sets contain the same number of indices of each kind and enjoy the same symmetries; consequently, they have the same number of essentially different values, N.

From this definition it is obvious that our matrix

$M = \| M^{a}_{b} \|$ is entirely characterized by N × N complex numbers $M^{k}_{l} \underset{df}{=} \pm M^{a}_{b}$ where $k, l = 1, 2, \dots, N$. (± to indicate that in the case when some skew symmetries are present, we have to pick up for M^{k}_{l} something with a definite sign). These are the independent components of the matrix. However, to the given component M^{k}_{l} there corresponds $p_k \cdot p_l$ combinations of the internal indices which all are leading to it (with accuracy up to the sign).

Now, because we want to construct the notion of the multiplication of the matrices as a tensorial notion (covariant with respect to the groups governing the internal indices) we define the multiplication rule as

$$M = M_1 M_2 \equiv \| M^{a}_{1s} M^{s}_{2b} \| = \| M^{a}_{1\,\{s'_1 \dots, s''_1 \dots\}} M_{2b}^{\{s'_1 \dots, s''_1 \dots\}} \| \tag{3.2}$$

where over all internal s-indices the standard summational convention applies. So defined multiplication is manifestly covariant with respect to G_1, G_2, ...

The usual multiplication of independent components of the type $\sum_{s=1}^{N} M_{1s}^{k} M_{2l}^{s}$ is - in general - a non-tensorial operation. Our rule (3.2) in terms of independent components may be understood as

$$M_l^k = \sum_{s=1}^{N} M_{1s}^{k} \, p_s \, M_{2l}^{s} \, . \tag{3.3}$$

Note that the ambiguity in signs associated with the choice of the complex numbers serving as the independent components of the matrix (present in the case of the skew symmetries) is here immaterial; in the summation the contra-independent values of the set $s, \overset{k}{\not{s}}$ always meets the corresponding co- $\overset{k}{\not{s}}$.

The unit N x N matrix we define as

$$\mathbf{1} \equiv \left\| E_b^a \right\| \equiv \left\| \delta^{\{A_1' \ldots A_1'' \ldots\}}_{\{B_1' \ldots B_1'' \ldots\}} \right\| \underset{\mathrm{df}}{=} \left\| \delta^{\{A_1'}_{B_1'} \ldots \delta^{A_1''}_{B_1''} \ldots \} \right\| \tag{3.4}$$

where $\delta^{A'}_{B'}$, etc. are the usual Kronecker δ's and { } stands for the symmetrizations of the type which is specific for the considered set of indices. The independent components of the matrix E_b^a clearly are

$$\overset{k}{E}{}^{\not{a}}_{\underset{\not{b}}{l}} = 0 \ \text{if} \ k \neq l, \ \overset{k}{E}{}^{a}_{\underset{\not{b}}{k}} = 1/p_k \ ,$$

with the proviso that in the case of skew symmetries we accept a suitable choice of signs for the independent components.

By the *trace* of M we will understand

$$T_r(\mathsf{M}) = \mathsf{M}^s_s = \mathsf{M}^{\{s_1' \dots s_1'' \dots\}}_{\{s_1' \dots s_1'' \dots\}} \tag{3.5}$$

where the summational convention applies over the internal indices. It is obvious that

$$T_r(\mathbf{1}) = E^s_s = N. \tag{3.6}$$

The unit matrix has the obvious fundamental property

$$\mathbf{1} \cdot \mathsf{M} = \mathsf{M} = \mathsf{M} \cdot \mathbf{1}, \tag{3.7}$$

which holds for any M.

From the technical point of view, it is convenient to consider simultaneouly with the set of indices a constructed from it the skew-symmetric *sets* of *sets* of indices

$$[a]_k = [a_1 a_2 \dots a_k], \quad k = 0, 1, \dots, N, \tag{3.8}$$

where the [ ] means the total skew-symmetrization with respect to the participating sets a_i; it does not affect their internal symmetries. For $k = 0$ our symbol becomes trivial; $[a]_0$ attached to a quantity means that it has only one component. The $[a]_1$ coincides with the set a itself. It is clear that the essentially different values of $[a]_k$ are of the form

$$[\alpha_1^{i_1}, \alpha_2^{i_2}, \ldots ; \alpha_k^{i_k}] \tag{3.9}$$

where $i_1 \ldots i_k$, are all different, and form a sequence picked up from the numbers $1,2,\ldots N$. Therefore, the set of sets of indices $[a]_k$ has exactly

$$N_k = \binom{N}{k} \tag{3.10}$$

essentially different values. It is obvious that the weight of (3.9) is just $k!\, P_{i_1}, P_{i_2} \ldots P_{i_k}$.

Now it is also technically convenient to consider together with our M_b^a (of the $N \times N$ type) the matrices:

$${}_kM \equiv \left\| M_{[b]_k}^{[a]_k} \right\| \in \mathcal{M}_k , \; k = 0,1,2,\ldots N \tag{3.11}$$

which are $N_k \times N_k$ matrices; by the symbol $\mathcal{M}_k$ we denote the set of such matrices. The $\mathcal{M}_0$ forms an ensemble of 1×1 matrices, i.e., scalars. The $\mathcal{M}_1$ is the ensemble of the $N \times N$ matrices studied.

Sometimes, when it is clear to which $\mathcal{M}_k$ the matrix ${}_kM$ belongs we will omit in the symbol ${}_kM$ the suffix k which appears on the left.

The unit matrix in $\mathcal{M}_k$ is obviously

$${}_k\mathbf{1} \equiv \left\| E_{[b]_k}^{[a]_k} \right\| = \left\| E_{b_1 \ldots b_k}^{[a_1 \ldots a_k]} \right\|, \; k = 0,1,\ldots, N . \tag{3.12}$$

For $k = 0$ we understand ${}_0\mathbf{1}$ just as the number 1. The trace of ${}_k\mathbf{1}$ is obviously

$$T_r\left({}_k\mathbf{1}\right) = E^{[s]_k}_{[s]_k} = N_k = \binom{N}{k} . \tag{3.13}$$

With the help of our unit matrices one defines the generalized Kronecker's symbols, δ's, as

$$\delta^{[a]_k}_{[b]_k} \underset{df}{=} k!\ E^{[a]_k}_{[b]_k} , \quad k = 0, 1, \ldots, N . \tag{3.14}$$

For $k = 0$ the Kronecker δ is just the number 1. The $\delta^a_b = \delta^{[a]_1}_{[b]_1}$ coincides with $E^{[a]_1}_{[b]_1} = E^a_b$. The advantage of these quantities lies in the fact that they may be represented as

$$\delta^{[a]_k}_{[b]_k} = \begin{vmatrix} \delta^{a_1}_{b_1} & \cdots & \delta^{a_k}_{b_1} \\ \vdots & & \vdots \\ \delta^{a_1}_{b_1} & \cdots & \delta^{a_k}_{b_k} \end{vmatrix} \tag{3.15}$$

i.e., as the usual determinant from single δ's.

Now the quantities of the type $T^{[a]_N}$ and $T_{[b]_N}$ because $N_N = \binom{N}{N} = 1$, have only one (essentially different) component. Therefore if $E^{[a]_N}$ and $E_{[a]_N}$

are the objects normalized so that

$$E^{[\alpha_1 \alpha_1' \dots \alpha_N]} = \sqrt{N!} = E_{[\alpha_1 \alpha_2 \dots \alpha_N]} \tag{3.16}$$

for a choice of numerical values of internal indices from sets of essentially different values as indicated, then necessarily $T^{[a]_N}$ and $T_{[b]_N}$ have to be proportional (with numerical factors) to $E^{[a]_N}$ and $E_{[a]_N}$. Because of the assumed skew symmetry which is always present, writing these objects we may omit the symbol [ ] $\left(E_{[a]_N} = E_{a_1 \dots a_N} \right)$ or we may divide the indices of these quantities into arbitrary – also skew – sub-sets of sets of indices

e.g., $E_{[a]_N} = E_{[a]_{i_1} \; [a]_{i_2} \dots [a]_{i_s}} \left(, \underline{i_1 + \dots + i_s =} \right)^N$. The normalization of these symbols is so chosen that the following is true

$$\binom{N}{k} E^{[a]_k [s]_{N-k}} E_{[b]_k [s]_{N-k}} = E^{[a]_k}_{[b]_k} \quad k = 0,1, \dots N \tag{3.17}$$

(These are the *partial traces*; note the important relation

$E^{[a]_N}_{[b]_N} = E^{[a]_N} E_{[b]_N}$). Also notice that

$$E^{[a_1 \dots a_{k-1} s]}_{[b_1 \dots b_{k-1} s]} = \frac{N-k+1}{k} E^{[a_1 \dots a_{k-1}]}_{[b_1 \dots b_{k-1}]} . \tag{3.18}$$

With the help of $E^{[a]_N}$ and $E_{[b]_N}$ one defines the generalized Levi-Civita symbols $\epsilon_{[a]_N} = 1/\sqrt{N!}\; E_{[a]_N}$ and $\epsilon^{[a]_N} = 1/\sqrt{N!}\; E^{[a]_N}$ (their normalization is $\epsilon_{[\alpha_1 \dots \alpha_n]} = 1 = \epsilon^{[\alpha_1 \dots \alpha_n]}$, in the same sense as in (3.16)).

Using these and the Kronecker δ's we may rewrite (3.17) and (3.18) as

$$\frac{1}{(N-k)!}\; \epsilon^{[a]_k [s]_{N-k}}\; \epsilon_{[b]_k [s]_{N-k}} = \delta^{[a]_k}_{[b]_k} \tag{3.19}$$

$$\delta^{[a]_{k-1}\, s}_{[b]_{k-1}\, s} = (n-k+1)\; \delta^{[a]_{k-1}}_{[b]_{k-1}} \tag{3.20}$$

For $k = 0$ (3.19) reduces to $\epsilon^{[s]_N}\; \epsilon_{[s]_N} = N!$; for $k = 1$ (3.20) reduces to $\delta^s_s = N$. These formulae are formally identical with those well-known from the tensor calculus. However, the fact that they hold in unchanged form for our sets of indices -although logically trivial- is not trivial from the technical point of view. It reduces algebraic manipulations with the quantities "wearing" sets of indices to those with which we are well familiar, moreover, it enables us all the time during these operations to keep the manifest covariancy.

IV. POWERS, SKEW POWERS, MINORS

The aim of this section is to extend the standard notions and results of the theory of determinants, minors, etc., to the case of our tensorial matrices.

Let $M = \| M^a_b \|$ be a $N \times N$ matrix from M_1. It generates in $\mathcal{M}_k$ the sequence of matrices

$$
{}_k M_{[q]} = \left\| M^{[a]_k}_{[q]} \; {}_{[b]_k} \right\| \underset{df}{=}
$$

$$
= \left\| \frac{N_q}{N_k} \; \delta^{[a_1}_{[b_1} \cdots \delta^{a_{k-q}}_{b_{k-q}} \; M^{a_{k-q+k}}_{b_{k-q+1}} \cdots M^{a_k]}_{b_k]} \right\| \tag{4.1}
$$

which are the skew symmetrized external products of the matrix M with itself and the unit matrix. By N_s we here understand $\binom{N}{s}$. These quantities are well-defined for $N \geq k \geq q \geq 0$. The index $[q]$ on the right indicates how many times this quantity contains the matrix M^a_b. For $q = 0$ we have

$$
{}_k M_{[0]} = (1/Nk) \; {}_k \mathbf{1} \; . \tag{4.2}
$$

The other extreme for $q = k$ defines the skew powers of the matrix:

$$M_{[k]} \underset{df}{=} {}_kM_{[k]} = \left\| M^{[a_1}_{b_1} \dots M^{a_k]}_{b_k} \right\| \in \mathcal{M}_k$$

(4.3)

Because these quantities from definition are objects from M_k we may omit in their symbols the k on the left. By $M_{[0]}$ we will understand just the number 1. The $M_{[1]}$ coincides with the matrix itself. Formally, the skew powers are defined for any k but for $k > N$ all of these identically vanish.

The skew powers enable us to introduce the concept of the rank of the matrix $M = \| M^a_b \|$; the matrix $M = {}_1M$ is said to be of the rank r if

$$N \geq k > r \rightarrow M_{[k]} \equiv 0, \quad r \geq k \geq 0 \rightarrow M_{[k]} \not\equiv 0 . \tag{4.4}$$

The traces of the skew powers define the fundamental sequence of invariants of the matrix:

$$\underset{[k]}{M} = T_r\left(M_{[k]}\right) = M^{[s_1}_{s_1} \dots M^{s_k]}_{s_k}$$

$$= \frac{1}{k!} \delta^{u_1 \dots u_k}_{v_1 \dots v_k} M^{v_1}_{u_1} \dots M^{v_k}_{u_k}, \quad k = 0,1,2,\dots, N .$$

(4.5)

By $\underset{[0]}{M}$ we understand just the number one, $\underset{[0]}{M} = 1$. For $\underset{[N]}{M}$ we also will use the alternative symbol

$$\text{Det} \left\| M^a_b \right\| \underset{df}{=} \underset{[N]}{M} . \tag{4.6}$$

Notice that for the matrix of rank r necessarily $N \geq k > r \rightarrow \underset{[k]}{M} = 0$.
The invariants with lower k may vanish as well as they may be different from zero.

Note that the normalization factor in (4.1) is so chosen that

$$T_r \left({}_k M_{[q]} \right) = \underset{[q]}{M} , \text{ for } N \geq k \geq q . \tag{4.7}$$

The skew powers are to be contrasted with the proper powers of the matrix, or simply, powers. These we define inductively as

$$M^{\circ} = \mathbf{1} , \; M^{p+1} = M \cdot M^p , \; p = 0,1,2,\ldots \quad . \tag{4.8}$$

Of course, with $M \in \mathcal{M}_1$ all its powers also are objects which belong to m_1

Note that the concept of power of the matrix may be extended along the same line for ${}_k M$ from $\mathcal{M}_k$:

$$ {}_k M^{\circ} = {}_k \mathbf{1} , \; {}_k M^{p+1} = {}_k M \cdot {}_k M^p , \; p = 0,1,2,\ldots \; . \tag{4.9}$$

The traces of the powers of $M = {}_{1}M$ form the sequence of invariants

$$T_r \left(M^p \right) = \overset{p}{M} \ , \ p = 0,1,2,\ldots \ . \tag{4.10}$$

For $p = 0$ we obviously have $\overset{o}{M} = N$; not all of these invariants are independent. But those with $p = 0,1,2,\ldots, N$ are independent. Consequently, we shall call $\overset{o}{M}, \overset{1}{M}, \ldots, \overset{N}{M}$ the *second fundamental sequence of invariants*.

The next important concept which we must introduce is that of *minors*.

$${}_k\mathfrak{m}_{[q]} = \left\| (1/k!q!) \, \delta^{[a]_k \, [s]_q}_{[b]_k \, [t]_q} \, M^{[t]_q}_{[q] \quad [s]_q} \right\| \in \mathcal{M}_k$$

$$N \geq k \geq 0 \ , \quad N \geq q \geq 0 \ , \quad k + q \leq N \ . \tag{4.11}$$

We have the obvious

$${}_k\mathfrak{m}_{[o]} = {}_k\mathbf{1} \ . \tag{4.12}$$

The minors in the extreme case $k + q = N$ we will call the proper minors, and they may be put in one to one correspondence with the skew powers:

$$M_{[k]} \leftrightarrow {}_k\mathfrak{m}_{[N-k]} \ . \tag{4.13}$$

Indeed, these are the algebraic complements of each other in the sense -

that the following holds:

$$ {}_{k}\mathfrak{m}_{[N-k]} \cdot M_{[k]} = \underset{[N]}{M} \cdot {}_{k}\mathbf{1} \,, \; k = 0,1,2,\ldots N \,. \tag{4.14} $$

(The equivalent of the Laplace's development of the determinant). Observe that

$$ T_r\left({}_{k}\mathfrak{m}_{[q]}\right) = \binom{N-q}{k} \underset{[q]}{M} \,. \tag{4.15} $$

(For the proper minors $T_r\left({}_{k}\mathfrak{m}_{[N-k]}\right) = \underset{[N-q]}{M}$. Also, from the formal point of view, setting $k = 0$ in (4.11), we conclude that

$$ {}_{0}\mathfrak{m}_{[q]} = \underset{[q]}{M} \,, \tag{4.16} $$

so that the fundamental sequence of invariants may be understood as a special case of minors.

Notice that the skew powers may be expressed through the proper minors as

$$ M_{[k]} = \left\| N_k \cdot {}_{N}E^{[a]_k \, [s]_{N-k}}_{[b]_k \, [t]_{N-k}} \; \mathfrak{m}^{[t]_{N-k}}_{[k]} \quad {}_{[s]_{N-k}} \right\| , \tag{4.17} $$

in obvious notation.

In order to learn more about the properties of skew powers, powers and minors, now we will study the so-called λ- matrices. Let

$$\mathsf{M}(\lambda) \underset{df}{=} \mathsf{M} - \lambda \cdot \mathbf{1} \tag{4.18}$$

where λ is a complex number. All of the previously considered identities are still true when we substitute $\mathsf{M} \to \mathsf{M}(\lambda)$; by comparing the coefficients of different powers of λ we will obtain a further sequence of useful identities.

First of all, substituting $\mathsf{M} \to \mathsf{M} - \mathbf{1} \cdot \lambda$ into (4.1) one easily finds

$$_k\mathsf{M}_{[q]}(\lambda) = \sum_{p=0}^{q} (-\lambda)^{q-p} \binom{N-p}{q-p} \, _k\mathsf{M}_{[p]} \, . \tag{4.19}$$

These general formulae for $k = q$ gives us the expression for $\mathsf{M}_{[q]}$:

$$\mathsf{M}_{[q]}(\lambda) = \sum_{p=0}^{q} (-\lambda)^{q-p} \binom{N-p}{q-p} \, _q\mathsf{M}_{[q]} \, . \tag{4.20}$$

Taking the traces of (4.19) or of (4.20) - because of (4.7) - we derive that

$$\underset{[q]}{\mathsf{M}}(\lambda) = \sum_{p=0}^{q} (-\lambda)^{q-p} \binom{N-p}{q-p} \underset{[p]}{\mathsf{M}} \quad q = 0,1,2,\ldots, N \, . \tag{4.21}$$

The last formula becomes specially important important for $q = N$ With the help of it we define the *characteristic polynomial* of the letter λ as

$$D_N(\lambda) \underset{df}{=} \mathrm{Det} \left\| \lambda \cdot \mathbf{1} - \mathsf{M} \right\| = (-)^N \underset{[N]}{\mathsf{M}}(\lambda) =$$

$$= (-1)^N \sum_{p=0}^{N} (-\lambda)^{N-p} \underset{[p]}{\mathsf{M}} \, . \tag{4.22}$$

(The λ^N here enters with the coefficient $\underset{[o]}{M} = 1$).

For the minors we find

$$ {}_k\mathfrak{m}_{[q]}(\lambda) = \sum_{p=0}^{q} (-\lambda)^{q-p} \binom{N-k-p}{q-p} {}_k\mathfrak{m}_{[p]} \tag{4.23}$$

which for the proper minors reduces to

$$ {}_k\mathfrak{m}_{[N-k]}(\lambda) = \sum_{p=0}^{N-k} (-\lambda)^{N-k-p} \, {}_k\mathfrak{m}_{[p]} \, . \tag{4.24}$$

Writing down (4.14) for the λ- matrices we derive

$$ \sum_{q=0}^{k} \sum_{p=0}^{N-k} (-\lambda)^{N-p-q} \binom{N-q}{k-q} {}_k\mathfrak{m}_{[p]} \cdot {}_kM_{[q]} $$

$$ = {}_k\mathbf{1} \cdot \sum_{n=0}^{N} (-\lambda)^{N-n} \underset{[n]}{M} \, . \tag{4.25}$$

Because the coefficients of powers of λ must agree, this splits into $(N+1)^2$ identities :

$$ \sum_{p=0}^{N-k} \sum_{q=0}^{k} \delta_{p+q,n} \binom{N-q}{k-q} {}_k\mathfrak{m}_{[p]} \cdot {}_kM_{[q]} = {}_k\mathbf{1} \underset{[n]}{M} \begin{cases} k = 0,1,\ldots N \\ n = 0,1,\ldots N \end{cases} . \tag{4.26}$$

These identities exhaust all the relevant information about the products of the type ${}_k\mathfrak{m}_{[p]} \cdot {}_k M_{[q]}$.

Particularly important are these identities for $k = 1$. Because ${}_1\mathfrak{m}_{[p]} \in \mathcal{M}_1$, similarly as the matrix $M = \underset{1}{M}$ itself, we may here omit the suffix 1 on the left, ${}_1\mathfrak{m}_{[p]} \equiv \mathcal{M}_{[p]}$. One easily sees that (4.26) for $k = 1$ may be written as

$$\mathfrak{m}_{[0]} = \underset{[0]}{M} \cdot \mathbf{1}$$

$$\mathfrak{m}_{[p]} + \mathfrak{m}_{[p-1]} \cdot M = \underset{[p]}{M} \cdot \mathbf{1} \ , \quad p = 1,2,\ldots,N-1$$

$$\mathfrak{m}_{[N-1]} \cdot M = \underset{[N]}{M} \cdot \mathbf{1} \ , \tag{4.27}$$

which are all $\mathfrak{m}_1$ equations. These equations immediately give us

$$\mathfrak{m}_{[p]} = \sum_{q=0}^{p} \left(-M\right)^{p-q} \underset{[q]}{M} \ , \quad p = 0,1,2,\ldots,N-1 \ , \tag{4.28}$$

an expression for the $N \times N$ minors through the proper powers and invariants. Using (4.28) written for $p = N-1$ in the last of (4.27) we obtain

$$\sum_{q=0}^{N} \left(-M\right)^{N-q} \underset{[q]}{M} = 0 \ . \tag{4.29}$$

This is equivalent to the statement that the characteristic polynomial taken from the matrix itself (the powers of the matrix understood as proper powers) vanishes:

$$\mathcal{D}_N \left(\mathsf{M} \right) = 0 \quad . \tag{4.30}$$

This is the Hamilton-Cayley equation for our tensorial matrices

$$\mathsf{M}^{\{A'_1 \ldots A''_1 \ldots\}}_{\qquad \{B'_1 \ldots B''_1 \ldots\}} \quad .$$

This statement holds in general, independent of the rank of the matrix. In the case of rank $k \leq r \leq N$ a similar but stronger statement is true. Define the incomplete characteristic polynomials as

$$\mathcal{D}_s (\lambda) = (-)^s \sum_{q=0}^{s} (-\lambda)^{s-q} \mathsf{M}_{[q]} \; , \; s = 0,1,2, \ldots, N \tag{4.31}$$

[For $s = N$ this becomes the full or complete characteristic polynomial]. Therefore, (4.28) may be rewritten as

$$m_{[p]} = (-)^p \; \mathcal{D}_p \left(\mathsf{M} \right) \; . \tag{4.32}$$

Now for the matrix of rank r we have $\mathsf{M}_{[r+1]} = 0$ which implies that

$\mathfrak{m}_{[r+1]} = 0$, $\underset{[r+1]}{\mathsf{M}} = 0$. Therefore, writing (4.27) with $p = r + 1$ we obtain

$$\mathsf{M} \cdot \mathcal{D}_r \left(\mathsf{M} \right) = 0 \qquad (4.33)$$

which is an equation of the $(r + 1)$'th order. The assumption $\mathsf{M}_{[r+1]} = 0$ while $\mathsf{M}_{[r]} \neq 0$ implies that $\underset{[r+1]}{\mathsf{M}} \ldots \underset{[N]}{\mathsf{M}}$ all vanish, so that (4.30) reduces to

$$\mathsf{M}^{N-r} \, \mathcal{D}_r \left(\mathsf{M} \right) = 0 \, . \qquad (4.34)$$

Therefore, for $r = N - 1$ (4.30) says the same thing as (4.33). However, for $N - r > 1$ the fact that the factor M^{N-r-1} may be removed from the identity (4.34) is not trivial.

One can remark that the more involved quantities of the theory ${}_k\mathsf{M}_{[q]}$ and ${}_k\mathfrak{m}_{[q]}$ as matrices from ${}^{\prime}\mathcal{M}_k$, also define their characteristic polynomials (of the order $N_k = \binom{N}{k}$) with invariant coefficients and – substituted to these polynomials – they annihilate them.

We now should like to introduce the concept of the inverse matrix. The matrix M is said to admit the existence of its inverse when there exists such matrix M^{-1} that

$$\mathsf{M} \cdot \mathsf{M}^{-1} = \mathbf{1} = \mathsf{M}^{-1} \cdot \mathsf{M} \, . \qquad (4.35)$$

The rank of the matrix which admits the inverse must necessarily be $r = N$.

Indeed, our definition of the determinant enjoys the property

$$\mathrm{Det} \left\| M_1 \cdot M_2 \right\| = \mathrm{Det} \left\| M_1 \right\| \cdot \mathrm{Det} \left\| M_2 \right\| . \tag{4.36}$$

Also: $\mathrm{Det} \left\| \mathbf{1} \right\| = 1$. Applying this rule in (4.35) one finds

$$\mathrm{Det} \left\| M \right\| \cdot \mathrm{Det} \left\| M^{-1} \right\| = 1 . \tag{4.37}$$

Therefore, for the rank $r < N$ we have a contradiction. But if $r = N$ so that $M_{[N]} = \mathrm{Det} \left\| M \right\| \neq 0$, the inverse is uniquely determined. Indeed, *the last of (4.27) yields*

$$M^{-1} = 1/ M_{[N]} \quad \mathfrak{m}_{[N-1]} . \tag{4.38}$$

Clearly in this case $\mathfrak{m}_{[N-1]} \neq 0$.

Now we would like to prove a simple theorem: if $P(\lambda)$ is a polynomial of the order n in the letter λ (with the coefficient of λ^n equal one), if $P(M) = 0$, then, for every λ such that $P(\lambda) \neq 0$ there exists the inverse matrix $[\mathbf{1} \cdot \lambda - M]^{-1}$.

Indeed, let

$$P_s(\lambda) = (-)^s \sum_{p=0}^{s} (-\lambda)^{s-p} \; P_{[p]} \quad ; \quad P_n(\lambda) \equiv P(\lambda) . \tag{4.39}$$

the $P_{[p]}$, $p = 0, 1, \ldots, n$ being some coefficients, $P_{[0]} \equiv 1$

Then one easily checks an algebraic identity:

$$\sum_{s=0}^{n-1} P_{n-1-s}(z)\, \lambda^s = \frac{P_n(\lambda) - P_n(z)}{\lambda - z} = \sum_{s=0}^{n-1} P_{n-1-s}(\lambda)\, z^s \,. \tag{4.40}$$

Therefore

$$(\lambda - z) \sum_{s=0}^{n-1} P_{n-1-s}(\lambda)\, z^s = P_n(\lambda) - P_n(z)\,. \tag{4.41}$$

Substituting into this identity $z \to M$ and making use of the assumption $P_n(\lambda) \neq 0$ we conclude that

$$\left[\mathbb{1} \cdot \lambda - M \right]^{-1} = \sum_{s=0}^{n-1} \frac{P_{n-1-s}(\lambda)}{P_n(\lambda)} \cdot M^s \,, \tag{4.42}$$

which proves our theorem, moreover, it provides us with the explicit construction of $\left[\mathbb{1} \cdot \lambda - M \right]^{-1}$.

V. THE MINIMAL POLYNOMIAL

Let

$$\overline{D}_s(\lambda) = (-)^s \sum_{p=0}^{s} (-\lambda)^{s-p} \underset{[p]}{\overline{M}}, \quad s = 0,1,2,\ldots \overline{N}; \quad \underset{[o]}{\overline{M}} \equiv 1 \tag{5.1}$$

be some polynomials with coefficients $\underset{[o]}{\overline{M}}, \ldots, \underset{[N]}{\overline{M}}$. The one with the highest order in this sequence, the $\overline{D}_{\overline{N}}(\lambda)$ is said to be the minimal polynomial if (1) $\overline{D}_{\overline{N}}\left(M\right) = 0$, (2) the number $\overline{N}$ is the lowest number for which a relation of this type can hold in a non-trivial manner.

The minimal polynomial enjoys a few important properties which are described by the following theorems:

Theorem I. The $\overline{D}_{\overline{N}}(\lambda)$ is unique. Therefore, all the family of $\overline{D}_s(\lambda)$ is uniquely determined. Indeed, if there exist two different minimal polynomials of the order $\overline{N}$, their difference would be a non-trivial polynomial of the order $< \overline{N}$ which can be annihilated by the substitution $\lambda \rightarrow M$.

Theorem II. The $\overline{D}_{\overline{N}}(\lambda)$ is a factor of any polynomial $P(\lambda)$ with the property $P\left(M\right) = 0$. Indeed, according to the definition of $\overline{N}$ the order of $P(\lambda)$ has to be $N' \geq \overline{N}$. Therefore, according to the rule about the division of polynomials

$$P(\lambda) = \overline{D}_{\overline{N}}(\lambda) \cdot S(\lambda) + R(\lambda), \tag{5.2}$$

where $S(\lambda)$ is of the order $N' - \overline{N}$, but $R(\lambda)$ is of the order $\overline{N} - 1$. Substituting $\lambda \to M$ we get $R(M) = 0$, which because its order is $< \bar{N}$ can be possible only when $R(\lambda) \equiv 0$.

Theorem III. The roots of the characteristic polynomial $D_N(\lambda)$ are roots of $\bar{D}_{\bar{N}}(\lambda)$ and inversely, the roots of $\bar{D}_{\bar{N}}(\lambda)$ are roots of $D_N(\lambda)$:

$$\bar{D}_{\bar{N}}(\lambda^*) = 0 \to D_N(\lambda^*) = 0\,,\; D_N(\lambda^*) = 0 \to \bar{D}_{\bar{N}}(\lambda^*) = 0. \tag{5.3}$$

Indeed, the truth of the first implication follows from theorem II according to which $D_N(\lambda)$ may be represented as

$$D_N(\lambda) = \bar{D}_{\bar{N}}(\lambda) \cdot \delta D(\lambda). \tag{5.4}$$

As far as the second implication is concerned, assume the opposite, i.e., that $D_N(\lambda^*) = 0$ but $\bar{D}_{\bar{N}}(\lambda^*) \neq 0$. Now taking in (4.42) as $P_n(\lambda^*)$ just $\bar{D}_{\bar{N}}(\lambda^*)$, we explicitly construct the well-defined matrix $[\mathbf{1} \cdot \lambda^* - M]^{-1}$. Now taking the determinant of the identity $[\mathbf{1} \cdot \lambda^* - M]^{-1} \cdot [\mathbf{1} \cdot \lambda^* - M] = \mathbf{1}$ and remembering that $D_N(\lambda^*) = \mathrm{Det}\,\|\mathbf{1} \cdot \lambda^* - M\|$ vanishes by assumption, we get $0 = 1$, i.e., the contradiction. Hence, the second implication is true

The characteristic polynomial (because complex numbers are algebraically closed) may be always represented as

$$D_N(\lambda) = \mathrm{Det}\left\| \lambda \cdot \mathbf{1} - M \right\| = (-)^N \sum_{p=0}^{N} (-\lambda)^{N-p} \underset{[p]}{M} = \prod_{i=1}^{N_0} \left(\lambda - M_i\right)^{n_i} \tag{5.5}$$

where

$$i \neq j \rightarrow M'_i - M'_j \neq 0, \; n_i \geq 1 \, , N = \sum_{i=1}^{N_0} n_i \, , \; N \geq N_0 \; . \tag{5.6}$$

The quantities M'_i , $i = 1, 2, \ldots, N_0$ – the different roots of $D_N(\lambda)$ – will be called *eigenvalues* of the matrix M. The integer exponents n_i are their *multiplicities*.

Now theorem III implies that if $D_N(\lambda)$ has the form of (5.5) then

$$\tilde{D}_{\bar{N}}(\lambda) = (-)^{\bar{N}} \sum_{\bar{p}=0}^{\bar{N}} (-\lambda)^{\bar{N}-p} \overline{M}_{[p]} = \prod_{i=1}^{N_0} \left(\lambda - M'_i\right)^{q_i} \tag{5.7}$$

where

$$n_i \geq q_i \geq 1 \, , \; \sum_{i=0}^{N_0} q_i = \bar{N} \, , \; N \geq \bar{N} \geq N_0 \; . \tag{5.8}$$

The $\overline{M}_{[p]}$ are uniquely defined invariants of the matrix because the $\tilde{D}_{\bar{N}}(\lambda)$ is itself unique. The quantities $\delta n_i = n_i - q_i$, the *algebraic defects* of the multiplicities are therefore arithmetic invariants of M. Note that $\delta D(\lambda)$ in (5.4) may be presented as

$$\delta D(\lambda) = \prod_{i=1}^{N_0} \left(\lambda - M'_i\right)^{\delta n_i} \; ; \; n_i - 1 \geq \delta n_i \geq 0 \, ,$$

$$\delta N = N - \bar{N} = \sum_{i=1}^{N_0} \delta n_i > 0 \, . \tag{5.9}$$

For the order of the minimal polynomial, $\bar{N}$, we always have the 'bound' $\bar{N} < N$. However, for the matrix of the rank r, $N - 1 \geq r \geq 1$ we have another stronger 'bound' for $\bar{N}$. Indeed, we found in the section IV that $M D_r(M) = 0$. Hence, according to theorem II we have $\lambda D_r(\lambda) = \bar{D}_{\bar{N}}(\lambda)\, S(\lambda)$ which is possible only for $\bar{N} \leq r + 1$. Moreover, because in this case $D_N(\lambda) \sim \lambda^{N-r} D_r(\lambda)$ one of the eigenvalues, say $\grave{M}_1$, must vanish; its multiplicity must obey $n_1 \geq N - r$. When additionally $\underset{[r]}{M} \neq 0$, we know more about this eigenvalue. Because $\lambda D_r(\lambda)$ here has only the single root $\lambda = 0$, it follows that $\bar{D}_{\bar{N}}(\lambda)$ also has only the single root $\lambda = 0$ so that $q_1 = 1$. However, $\underset{[r]}{M} \neq 0$ implies that $n_1 = N - r$, so that $\delta n_1 = N - r - 1$.

Now comparing the coefficients of λ in (5.5) we find

$$\underset{[p]}{M} = \underset{[p]}{\Phi}\left(\grave{M}_1, \ldots, \grave{M}_{N_0}\right) =$$

$$= \frac{(-1)^p}{2\pi i} \oint_{C_0} \frac{d\lambda}{\lambda^{N-p+1}} \prod_{i=1}^{N_0} (\lambda - \grave{M}_i)^{n_i} \quad p = 0, 1, \ldots, N, \tag{5.10}$$

so that the invariants $\underset{[p]}{M}$ are uniquely defined functions of eigenvalues. Among these, particularly simple are

$$\underset{[1]}{M} = \underset{[1]}{\Phi}\left(M'_1, \ldots, M'_{N_0}\right) = \sum_{i=1}^{N_0} n_i M'_i \tag{5.11}$$

$$\underset{[N]}{M} = \underset{[N]}{\Phi}\left(M_1',\ldots, M_{N_0}'\right) = \prod_{i=0}^{N_0} M_i'^{\,n_i}$$

Now let $F(z) = a \prod_{j=1}^{n} (z_j - z)$ be any polynomial of z (the z_i do not need to be different). Substituting $z \to M$ we have a well-defined matrix $F = F(M)$. Now because the determinant of the product of matrices is eaual to the product of the determinants we have

$$\text{Det}\left\| F(M) \right\| = \text{Det}\left\| a \prod_{j=1}^{n} (z_j - M) \right\| =$$

$$= a^N \prod_{j=1}^{n} \text{Det}\left\| z_j - M \right\| = a^N \prod_{j=1}^{n} D_N(z_j) = a^N \prod_{j=1}^{n} \prod_{i=1}^{N_0} \left(z_j - M_i'\right)^{n_i} =$$

$$= \prod_{i=1}^{N_0} \left(a \prod_{j=1}^{n} \left(z_j - M_i'\right)\right)^{n_i} = \prod_{i=1}^{N_0} \left[F\left(M_i'\right)\right]^{n_i} \tag{5.12}$$

Replacing in the derived identity $F(M)$ by $\mathbf{1}\lambda - F(M)$ we obtain

$$\text{Det}\left\| \mathbf{1}\lambda - F(M) \right\| = (-)^N \sum_{p=0}^{N} (-\lambda)^{N-p} \underset{[p]}{F} = \prod_{i=1}^{N_0} (\lambda - F(M_i'))^{n_i}$$

This relation shows that the eigenvalues of $F(M)$ are simply $F(M_i^{`})$; if it happens that $i \neq j \rightarrow F(M_i^{`}) \neq F(M_j^{`})$ their multiplicities are the same as those of the original matrix. Moreover, comparing the coefficients of powers of λ in (5.13) we conclude that the invariants of the matrix $F = F(M)$, $\underset{[p]}{F}$, are given as

$$\underset{[p]}{F} = \underset{[p]}{\Phi}\left(F(M_1^{`}), \ldots, F(M_{N_0}^{`})\right) \tag{5.14}$$

with the same functions $\underset{[p]}{\Phi}$ as those defined in (5.19). In particular, observing that $\underset{[1]}{M} = M_s^s = \overset{1}{M} = \mathrm{Tr}(M)$ we conclude that

$$\mathrm{Tr}\left[F(M)\right] = \sum_{i=1}^{N_0} n_i \, F(M_i^{`}) \,. \tag{5.15}$$

This formula will be of importance later.

Note that we have proven (5.13) and consequently (5.14) and (5.15) only for the polynomials $F(\lambda)$. But these formulae also happen to be true for more general functions on matrices $F(M)$, which we will study later. For the moment observe this: if $G(z)$ is a polynomial with roots z_i for which the inverse matrix $\{z_i - M^{-1}$ exists, we may understand - if $G(z) = b \prod_{j=1}^{m} (z_j - z)$ - the $G^{-1}(M)$ as the product

$$G^{-1}(M) = \frac{1}{b} \prod_{j=1}^{m} \left[z_j - M\right]^{-1} \tag{5.16}$$

of corresponding inverse matrices. Therefore, the rational functions $F(M)/G(M)$ of M are well-defined. It is obvious that what was crucial in the derivation of (5.13) - (5.15) was the product representation of the $F(M)$ studied. This clearly remains true when $F(\lambda)$ happens to be a ratio of two polynomials. Therefore, (5.13) - (5.15) are also valid for rational functions. Applying a limiting process in the numerator and the denominator so that $F(z)$ and $G(z)$ tend to infinite products or series one concludes that the formulae considered also remain valid when $F(\lambda)$ happens to be a meromorphic function such that $F(M)$ makes sense. We will return to these problems later.

VI. RELATIONSHIPS BETWEEN INVARIANTS

In this section we would like to study the relationships between various kinds of invariants which one can construct from the matrix $\mathsf{M} \in \mathcal{M}_1$.

There are two fundamental kinds of invariants

$$\underset{[p]}{\mathsf{M}} = \mathrm{Tr}\left(\mathsf{M}_{[p]}\right) , \; p = 0,1,\ldots,N \; ; \tag{6.1}$$

$$\overset{p}{\mathsf{M}} = \mathrm{Tr}\left(\mathsf{M}^{p}\right) , \; p = 0,1,2,\ldots \quad . \tag{6.1}$$

The question arises: how are these related. To answer this question, observe first that (5.15) in particular implies

$$\overset{p}{\mathsf{M}} = \sum_{i=1}^{N_0} n_i \, \mathsf{M}_i^{\prime\, p} \quad p = 0,1,2,\ldots \quad . \tag{6.2}$$

Now introduce the polynomials constructed from the incomplete characteristic polynomials as

$$\psi_s(z) = z^s\left(D_N\right)\frac{1}{z} = \sum_{q=0}^{s} (-z)^q \underset{[q]}{\mathsf{M}} \; , \; s = 0,1,\ldots N \, . \tag{6.3}$$

(Observe that $\psi_s(0) = 1$). Now using the product representation $D_N(z)$ from (5.5) we conclude that

$$\psi_N(z) = z^N D_N\left(\frac{1}{z}\right) = \sum_{q=0}^{N} (-z)^q \underset{[q]}{M} = \prod_{i=1}^{N_0} \left(1 - z M_i^{`}\right)^{n_i} . \tag{6.4}$$

Taking the logarithm of this relation we derive

$$\ln\left(\psi_N(z)\right) = \sum_{i=1}^{N_0} n_i \ln\left(1 - z M_i^{`}\right) . \tag{6.5}$$

Therefore, using on the right side of this equation, the uniformly convergent development of $\ln\left(1 - z M_i^{`}\right)$ [for $|z| < \mathrm{Min}\left\{ |M_1^{`}|^{-1}, \ldots, |M_{N_0}^{`}|^{-1} \right\}$] we obtain

$$\ln\left(\psi_N(z)\right) = -\sum_{i=1}^{N_0} n_i \sum_{p=1}^{\infty} \frac{z^p}{p} M_i^{`p} = -\sum_{p=1}^{\infty} \frac{z^p}{p} \left( \sum_{i=1}^{N_0} n_i M_i^{`p} \right). \tag{6.6}$$

Here using (6.2) we derive an identity which is crucial in this section

$$\ln\left\{ \sum_{q=0}^{N} (-z)^q \underset{[q]}{M} \right\} = -\sum_{p=1}^{\infty} \frac{z^p}{p} \overset{p}{M} \text{, for:}$$

$$|z| < R \underset{df}{=} \operatorname{Min}\left\{ \left|M_1^{\backprime}\right|^{-1}, \ldots, \left|M_{N_0}^{\backprime}\right|^{-1} \right\} . \tag{6.7}$$

This identity, valid in the radius of convergence, R for any z, already does not explicitly contain either $M_i^{\backprime}$'s or n_i's.

It enables us to immediately write down a sort of Rodriguez' formula in which we express one type of invariant through the other

$$\underset{[p]}{M} = \frac{(-1)^p}{p!}\left(\frac{d}{dz}\right)^p \exp\left[- \sum_{s=1}^{p,N,\infty} \frac{z^s}{s} \overset{s}{M}\right]\Bigg|_{z=0} ; \ p = 0,1,\ldots,N . \tag{6.8}$$

Here the upper limit of summation may be p or one can go with s up to N or up to ∞, as is convenient.

$$\overset{p}{M} = - \frac{1}{(p-1)!}\left(\frac{d}{dz}\right)^p \ln\left[\sum_{q=0}^{p,N} (-z)^q \underset{[q]}{M}\right]\Bigg|_{z=0} , \ p = 1,2,\ldots . \tag{6.9}$$

The upper limit of summation for $p < N$ may be p; for $p > N$ it must be kept as N.

These formulae, although looking very innocent, solve in a compact form a rather difficult problem: expressing the determinant $\underset{[N]}{M}$ through power invariants $\overset{p}{M}$ obviously is always possible; but for the determinants of higher order it is rather involved when we just start from the definition of the determinant.

A few additional coments: (6.7) in the form

$$\sum_{q=0}^{N} (-z)^q \underset{[q]}{M} = \exp\left[- \sum_{p=1}^{\infty} \frac{z^p}{p} \overset{p}{M} \right] \tag{6.10}$$

implies not only (6.8) but also

$$\left(\frac{d}{dz}\right)^p \exp\left[- \sum_{s=1}^{p} \frac{z^s}{s} \overset{s}{M} \right] \Bigg|_{z=0} = 0 \text{ for } p = N+1, N+2, \ldots \quad . \tag{6.11}$$

This - by applying the Leibnitz rule of differentiation - reduces to

$$p+1 > N \rightarrow \overset{p+1}{M} = \frac{1}{p!} \left(\frac{d}{dz}\right)^{p+1} \exp\left[- \sum_{s=1}^{p} \frac{z^s}{s} \overset{s}{M} \right] \Bigg|_{z=0} \tag{6.12}$$

This formula gives us the step-by-step rule of how to express $\overset{p}{M}$ for the arbitrary $p > N$ through the independent invariants $\overset{1}{M}, \ldots, \overset{N}{M}$. Indeed, (6.12) for $p = N$ gives us $\overset{N+1}{M}$ through $\overset{1}{M} \ldots \overset{N}{M}$. Knowing $\overset{1}{M} \ldots \overset{N+1}{M}$ writing (6.12) for $p = N+1$ we derive $\overset{N+2}{M}$, etc.

Now expression (6.9) suffers one technical disadvantage: it fails to cover the case $p = 0$ in which we should have $\overset{0}{M} = N$. We claim that the uniform expression for $\overset{p}{M}$ through $\underset{[1]}{M} \ldots \underset{[N]}{M}$, valid for all p's, including $p = 0$, may be written as

$$\overset{p}{M} = -\frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^p} \frac{d}{dz} \ln \left\{ D_N\left(\frac{1}{z}\right) \right\}, \quad p = 0,1,2,\ldots \tag{6.13}$$

where $D_N\left(\frac{1}{z}\right) = (-)^N \sum_{q=0}^{N} (-z)^{q-N} \underset{[q]}{M}$. One can verify this directly by computing the integral for $p = 0$ (which happens to be N) and for $p \geq 1$ rewriting $\ln D_N\left(\frac{1}{z}\right)$ as $\ln\left(z^{-N} \psi_N(z)\right) = -N \ln z + \ln \psi_N(z)$ which proves that for $p > 1$ (6.13) reduces to (6.9) . The other method of checking it consists of the substitution of $1/z = \xi$ in (6.13) (remembering the necessity of changing the direction of integration) . That way one gets

$$\overset{p}{M} = \frac{1}{2\pi i} \oint_{C_\infty} \xi^p \frac{d}{d\xi} \left(\ln D_N(\xi) \right) d\xi =$$

$$= \frac{1}{2\pi i} \oint_{C_\infty} d\xi \, \xi^p \sum_{i=1}^{N_0} \frac{n_i}{\xi - M'} = \sum_{i=1}^{N_0} n_i \, M_i'^{\,p} \, , \tag{6.14}$$

as expected.

In the further considerations we will often use the incomplete characteristic polynomials $D_s(\lambda)$ [see (4.31) for the definition] , or the generated by these polynomials $\psi_s(\lambda)$ [see (6.4)] . These, according to their definitions, are determined by the invariants $\underset{[p]}{M}$, from the first fundamental sequence, which serve as their coefficients. We know, however, that $\underset{[p]}{M}$ may be uniquely expressed by

$\overset{1}{M} \ldots \overset{N}{M}$, from the second fundamental sequence. It is of interest to express the polynomials $\psi_s(\lambda)$, $D_N(\lambda)$ explicitly through the invariants from the second fundamental sequence.

In (6.4) using the expressions (6.8) for $M_{[p]}$ we have

$$\psi_s(\lambda) = \sum_{p=0}^{s} (-\lambda)^p \, M_{[p]} =$$

$$= \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{d} \sum_{p=0}^{s} \left(\frac{\lambda}{z}\right)^p \exp\left[- \sum_{q=1}^{s} \frac{z^q}{q} \overset{q}{M}\right] =$$

$$= \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{\lambda - z} \left(\left[\frac{\lambda}{z}\right]^{s+1} - 1\right) \exp\left[- \sum_{q=1}^{s} \frac{z^q}{q} \overset{q}{M}\right] \tag{6.15}$$

If λ is located outside C_0 the second term in the integral, as an analytic, does not contribute. In that which remains we set $z = \lambda\xi$ obtaining

$$\psi_s(\lambda) = \frac{1}{2\pi i} \oint_{C_0} \frac{d\xi}{1-\xi} \frac{1}{\xi^{s+1}} \exp\left[- \sum_{q=1}^{s} \frac{\xi^q}{q} \overset{q}{M} \lambda^q\right] \quad ; \tag{6.16}$$

it is obvious that in this integral we may replace $1/(1-\xi)$ by

$$\exp \quad -\ln(1-\xi) \quad \rightarrow \exp\left[\sum_{q=1}^{s} \frac{\xi^q}{q}\right] \quad \text{so that}$$

$$\psi_s(\lambda) = \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^{s+1}} \exp\left[\sum_{q=1}^{s} \frac{z^q}{q}\left(1-\lambda^q \overset{q}{M}\right)\right] \tag{6.17}$$

and, consequently

$$\psi_s(\lambda) = \frac{1}{s!}\left(\frac{d}{dz}\right)^s \exp\left[\sum_{p=1}^{s} \frac{z^p}{p}\left(1-\lambda^p \overset{p}{M}\right)\right]\Bigg|_{z=0} \qquad s = 0,1,\ldots,N. \tag{6.18}$$

Which is an explicit expression for the polynomials $\psi_s(\lambda)$ in terms of the invariants $\overset{p}{M}$. Because $D_s(\lambda) = \lambda^s \psi_s\left(\frac{1}{\lambda}\right)$, using (6.18) we easily find

$$D_s(\lambda) = \frac{1}{s!}\left(\frac{d}{dz}\right)^s \exp\left[\sum_{p=1}^{s} \frac{z^p}{p}\left(\lambda^p - \overset{p}{M}\right]\right|_{z=0} \qquad s = 0,1,\ldots,N. \tag{6.19}$$

It is interesting to observe that with the help of this formula one may easily obtain a simple identity

$$\sum_{p=0}^{N-1} D_{N-1-p}(\lambda)\, w^p = \frac{1}{(N-1)!}\left(\frac{d}{dz}\right)^{N-1} \exp\left[\sum_{p=1}^{N-1} \frac{z^p}{p} \quad \lambda^p + w^p - \overset{p}{M})\right] \tag{6.20}$$

for arbitrary λ and w; it is obviously symmetric in them.

One can add the remark that the relationships between $\underset{[p]}{M}$'s and $\overset{p}{M}$ may be also approached from another angle. Indeed, by taking the trace of (4.28) [remembering (4.15)] one finds

$$q \cdot \underset{[q]}{M} + \sum_{p=0}^{q-1} (-)^{q-p} \overset{q-p}{M} \underset{[p]}{M} = 0 , \quad q = 1,2,\ldots,N-1 . \tag{6.21}$$

On the other hand, multiplying $D_N(M) = 0$ by M^l and taking the trace one has

$$\sum_{k=0}^{N} \overset{k+l}{M} (-)^k \underset{[N-k]}{M} = 0 , \quad l = 0,1,2,\ldots \quad . \tag{6.22}$$

These algebraic relations of invariants of both types form a chain of equations which enable us - by a sequence of successive steps - to find $\overset{p}{M}$ - of any desired order - as expressed through $\underset{[p]}{M}$, $p = 1,\ldots,N$. The inverse problem entirely solves the system of relations (6.21), which enables us to express step-by-step any $\underset{[q]}{M}$ through $\overset{1}{M} \ldots \overset{q}{M}$. The relations (6.8) - (6.9) are essentially equivalent to (6.21) and (6.22) . However, paying the price of the higher order differentiations we have the compact expressions which do not require any step-by-step treatment.

Now, we should like to introduce a third class of invariants which has certain advantages in comparison with the $\underset{[p]}{M}$'s or $\overset{p}{M}$'s; they enable us to approach the problem of the algebraic criteria for the existence of N_0 of different eigenvalues. We define the invariants of the third kind as

$$\Delta_s = \mathrm{Det}\left\| \overset{k+l}{M} \right\| , \quad k, l = 0, 1, \ldots, s-1 , \quad s = 1, 2, \ldots \tag{6.23}$$

where the matrix $\left\| \overset{k+l}{M} \right\|$ and the determinant are to be understood in the standard sense.

In order to see the meaning of these invariants, write them explicitly as

$$\Delta_s = \mathrm{Det}\left\| \overset{k+l}{M} \right\| = \frac{1}{s} \, \epsilon_{k_0 \cdots k_{s-1}} \, \epsilon_{l_0 \cdots l_{s-1}} \overset{k_0 + l_0}{M} \cdots \overset{k_{s-1} + l_{s-1}}{M}$$

$$k, = 0, 1, \ldots, s-1 . \tag{6.24}$$

But taking into account that $\overset{p}{M} = \sum_{i=1}^{N_0} n_i \, M_i'^{\,p}$, we may rewrite it as

$$\Delta_s = \sum_{i_0 = 1}^{N_0} \cdots \sum_{i_{s-1} = 1}^{N_0} n_{i_0} \cdots n_{i_{s-1}} \frac{1}{s!} \, \epsilon_{k_0 \cdots k_{s-1}} \, \epsilon_{l_0 \cdots l_{s-1}} \, M_{i_0}'^{\,k_0 + l_0} \cdots M_{i_{s-1}}'^{\,k_{s-1} + l_{s-1}} \tag{6.25}$$

Now recall the definition of the Van der Mond's determinant

$$V(x_1, \ldots, x_n) \underset{df}{=} \begin{vmatrix} x_1^0 & \ldots & x_n^0 \\ \vdots & & \vdots \\ x_1^{n-1} & & x_n^{n-1} \end{vmatrix} \tag{6.26}$$

and its fundamental property

$$V(x_1, \ldots, x_n) = \prod_{i=2}^{n} (x_i - x_1) \cdot V(x_2, \ldots, x_n) = \prod_{k=2}^{n} \prod_{i=k}^{n} (x_i - x_{k-1}) \tag{6.27}$$

With this definition of V as the function of many variables, it is obvious that (6.25) reduces to

$$\Delta_s = \frac{1}{s!} \sum_{i_0=1}^{N_0} \cdots \sum_{i_{s-1}=0}^{N_0} n_{i_0} \ldots n_{i_{s-1}} V^2 \left(M'_{i_0}, \ldots M'_{i_{s-1}} \right) \tag{6.28}$$

This form of Δ_s, along with the fact that V is skew-symmetric with respect to permutations of the variables, immediately implies that

$$s > N_0 \rightarrow \Delta_s = 0, \ \Delta_{N_0} = n_1 n_2 \ldots n_{N_0} \ V^2 \left(M'_1, \ldots, M'_{N_0} \right) \neq 0 . \tag{6.29}$$

This implies that the highest non-vanishing Δ_s defines the number of different eigenvalues. In the case of N different eigenvalues $\Delta_N = V^2\left(\overset{\prime}{M}_1, \ldots, \overset{\prime}{M}_N\right) \neq 0$ and $\Delta_{N+s} = 0, \ s = 1,2,\ldots$.

Another interesting feature of the invariants Δ_s is that in the case of *real* eigenvalues all of these invariants are non-negative. One can mention that Δ_s may be represented as

$$\Delta_s = \frac{1}{s}\left(-\frac{1}{2\pi i}\right)^s \oint_{C_0^1} \cdots \oint_{C_0^s} d\ \ln\left\{ D_N(z_1^{-1})\right\} \ldots d\ \ln\left\{ D_N(z_s^{-1})\right\} \qquad V^2(z_1^{-1}, \ldots, z_s^{-1}), \tag{6.30}$$

which in principle expresses them through the $\underset{[p]}{M}$'s. But these formulas are of little practical value. The most compact expression for Δ_s is that which serves as its definition, the expression through the $\overset{p}{M}$'s.

VII. THE MATRICES AS TRANSFORMATIONS

Up to now we did not need to introduce the notion of *vectors* upon which our matrices may act thus changing them into other vectors. There are, however, certain aspects in the analysis of matrices which will become clearer when one uses this approach.

The quantities (complex) of the type v^a, v_a 'wearing' *one* co- or contra-set of indices $a = \{ A_1' \dots A_1'' \dots \}$ may be understood as the elements of linear vector spaves $V^\bullet$, $V_\bullet$; (the contra- and co- vector spaces); we will refer to them by the word 'vectors'. All of the standard theory of the linear vector spaces applies, of course, to the case of our vectors 'wearing' *sets* of indices.

Without entering into any details, we will recall a few definitions and consequences which follow from them.

The vectors $\underset{i}{v}^a \in V^\bullet$, $i = 1,2,\dots,n$ are said to be linearly independent if

$$\overset{i}{x}\, \underset{i}{v}^a = 0 \rightarrow \overset{i}{x} = 0 \tag{7.1}$$

($\overset{i}{x}$ are numbers; over i the summational convention is assumed).

The necessary and sufficient condition for the linear independence reads

$$\underset{[i_1}{v}^{a_1} \dots \underset{i_n]}{v}^{a_n} \neq 0 \tag{7.2}$$

[The 'volume of a parallelipiped' spanned on $\underset{1}{v}^a \dots \underset{n}{v}^a$ has to be $\neq 0$].

The necessity of (7.2) is self-evident; the sufficiency may easily be demonstrated. Namely, because i takes on values only between $1 \ldots n$,

$$\underset{[i_1}{v}^{a_1} \ldots \underset{i_n]}{v}^{a_n} = \epsilon_{i_1 \ldots i_n} \, v^{a_1 \ldots a_n}$$

$$v^{a_1 \ldots a_n} = \frac{1}{n!} \, \epsilon^{i_1 \ldots i_n} \underset{i_1}{v}^{a_1} \ldots \underset{i_n}{v}^{a_n} \quad . \tag{7.3}$$

Therefore

$$\frac{1}{(n-1)!} \quad \epsilon^{ii_2 \ldots i_n} \underset{[j}{v}^{a_1} \underset{i_2}{v}^{a_2} \ldots \underset{i_n]}{v}^{a_n} = v^{a_1 \ldots a_n} \, \delta^i_j \tag{7.4}$$

This, contracted with $\overset{j}{x}$ – (when it is assumed that $\overset{j}{x} \underset{j}{v}^a = 0$) – yields

$$v^{a_1 \ldots a_n} \overset{i}{x} = 0 \rightarrow \overset{i}{x} = 0 \text{ when } v^{a_1 \ldots a_n} \neq 0 \, . \tag{7.5}$$

Therefore, (7.2) is necessary and sufficient for the implication of (7.1).

The obvious

$$\underset{[i_1}{v}^{a_1} \ldots \underset{i_n]}{v}^{a_n} = \underset{i_1}{v}^{[a_1} \ldots \underset{i_n}{v}^{a_n]} \tag{7.6}$$

-- recalling that from the definition of the set *a* there does not exist a quantity totally skew in more than N sets $a_1 \ldots a_n$ – implies that the maximal number of linearly independent vectors in our vector spaces is N . For this reason they shall be denoted as V^N , V_N . From now on all that is said about contra-vectors will apply 'mutatis mutandum' to co-vectors. Any N linearly independent vectors form a *basis*. Let e^a, $i = 1, \ldots, N$ form the basis. The necessary and sufficient condition for it can be seen to be

$$\underset{(N)}{e} = \frac{1}{N!} \epsilon^{i_1 \ldots i_n} \epsilon_{a_1 \ldots a_N} \underset{i_1}{e}^{a_1} \ldots \underset{i_N}{e}^{a_N} \neq 0 \tag{7.7}$$

A contra-basis $\underset{i}{e}^a \in V^N$ uniquely defines the *inverse* basis $\overset{i}{e}_a$ in V_N according to

$$\underset{i}{e}^a \overset{j}{e}_a = \delta^j_i \longleftrightarrow \overset{i}{e}_b \underset{i}{e}^a = \delta^a_b \quad . \tag{7.8}$$

The explicit expressions for $\overset{i}{e}_a$ in terms of $\underset{i}{e}^a$ are

$$\overset{i}{e}_a = \frac{1}{(N-1)!} \underset{(N)}{[e]}^{-1} \epsilon^{i i_2 \ldots i_N} \epsilon_{a a_2 \ldots a_N} \underset{i_2}{e}^{a_1} \ldots \underset{i_N}{e}^{a_N} \quad . \tag{7.9}$$

With the help of these concepts some new information concerning our

N x N tensorial matrices may be obtained. Using (7.8) we may write

$$\underset{b}{M}^{a} = \underset{s}{M}^{a} \underset{i}{e}^{s} \overset{i}{e}_{b} = \underset{i}{M}^{a} \overset{i}{e}_{b} \; ; \; \underset{i}{M}^{a} \underset{df}{=} \underset{s}{M}^{a} \underset{i}{e}^{s} \quad . \tag{7.10}$$

Therefore, the skew powers of M may be written as

$$\underset{b_1}{M}^{[a_1} \cdots \underset{b_q}{M}^{a_q]} = \underset{i_1}{M}^{[a_1} \cdots \underset{i_q}{M}^{a_q]} \overset{i_1}{e}_{b_1} \cdots \overset{i_q}{e}_{b_q} \quad . \tag{7.11}$$

This simple fact has the implication that if the matrix is of the rank $r = N$, all vectors $\underset{i}{M}^{a} = \underset{s}{M}^{a} \underset{i}{e}^{s}$ are linearly independent; the same may be expressed by saying that the matrix of the rank N transforms the basis into a new basis.

For the matrix of the rank r, taking (7.11) for $q = r + 1$ and using the first of (7.8) we conclude that

$$\underset{[i_1}{M}^{a_1} \cdots \underset{i_{r+1}]}{M}^{a_{r+1}} = 0 \tag{7.12}$$

so that any of the $r + 1$ vectors $\underset{i}{M}^{a}$'s are linearly dependent. Similarly one shows from $M_{[r]} \neq 0$ that among the $\underset{i}{M}^{a}$'s there certainly exist r linearly independent vectors. These facts, together with the representation of the matrix as (7.19), enable us to prove that the matrix of the rank r may always be represented as

$$M^a_b = \sum_{i=1}^{r} \underset{i}{\xi}^a \overset{i}{\eta}_b \tag{7.13}$$

where ξ's and η's are linearly independent (respectively in V^N and V_N).

In particular, any matrix of the rank $r = 1$ may be represented as $M^a_b = \xi^a \eta_b$; a matrix of this form is said to be dyadic. Therefore, the matrix of the rank r is a sum of r dyadic matrices.

Suppose that the matrix M is given in the form (7.13). The r linearly independent $\underset{i}{\xi}^a$ always may be completed by some $\underset{r+1}{\xi}^a, \dots \underset{N}{\xi}^a$ to a basis in V^N. This basis defines the corresponding inverse basis $\overset{i}{\xi}_a \in V^N$. The same may be repeated with respect to the $\overset{i}{\eta}_a$'s; we complete the set of η's by adding some $N - r$ vectors to a basis in V^N, and this basis defines the corresponding inverse basis $\underset{i}{\eta}^a \in V^N$. On these vectors one can span the following vectorial sub-spaces:

$$V^{(r)} \ni \sum_{i=1}^{r} \overset{i}{x} \underset{i}{\xi}^a, \quad V^{(N-r)} \ni \sum_{i=r+1}^{N} \overset{i}{x} \underset{i}{\xi}^a, \quad V^N = V^{(r)} \oplus V^{(N-r)}$$

$$\bar{V}^{(r)} \ni \sum_{i=1}^{r} \overset{i}{x} \underset{i}{\eta}^a, \quad \bar{V}^{(N-r)} \ni \sum_{i=r+1}^{N} \overset{i}{x} \underset{i}{\eta}^a, \quad V^N = \bar{V}^{(r)} \oplus \bar{V}^{(N-r)}$$

$$V_{(r)} \ni \sum_{i=1}^{r} x_i \, \eta^{i}_{a} \,, \quad V_{(N-r)} \ni \sum_{i=r+1}^{N} x_i \, \eta^{i}_{a} \,, \quad V_N = V_{(r)} \oplus V_{(N-r)}$$

$$\bar{V}_{(r)} \ni \sum_{i=1}^{r} x_i \, \xi^{i}_{a} \,, \quad \bar{V}_{(N-r)} \quad \sum_{i=r+1}^{N} x_i \, \xi^{i}_{a} \,, \quad \bar{V}_N = \bar{V}_{(r)} \oplus \bar{V}_{(N-r)} \,.$$

(7.14)

Now, the linear relations of the types

$$\text{(i)} \;\; y^a = M^a_s \, x^s \,, \quad \text{(ii)} \;\; y_b = x_s \, M^s_b \tag{7.15}$$

may be understood as the linear transformations in V^N or V_N respectivelly.

Consider first the case with the rank $(M) = N$. In this case, because M transforms a basis into another basis, (7.15) may be understood as the mappings

$$V^N \to V^N \,, \quad V_N \to V_N \,, \tag{7.16}$$

of our vector spaces into themselves ; we may express this symbolically as

$$V^N = M \, V^N \,, \quad V_N = V_N \, M \tag{7.17}$$

then to *any* y^a or y_a there exists a *unique* x^a or x_a:

$$x^a = M^{-1a}_{\ \ s} \, y^s \quad , \quad x_b = y_s \, M^{-1s}_{\ \ b} \quad . \tag{7.18}$$

Consider now the case $1 < r < N$. It is obvious that in this case the transformation (7.15) corresponds to the mappings:

$$V^{(r)} \to V^{(r)} \quad , \quad V^{(N-r)} \to 0$$

$$V_{(r)} \to V_{(r)} \quad , \quad V_{(N-r)} \to 0 \tag{7.19}$$

or, more symbolically:

$$v^{(r)} = M \, \bar{V}^{(r)} \quad , \quad 0 = M \, \bar{V}^{(N-r)}$$

$$v_{(r)} = \bar{V}_{(r)} \, M \quad , \quad 0 = \bar{V}_{(N-r)} \, M \; . \tag{7.20}$$

The sub-spaces $\bar{V}^{(N-r)}$ and $\bar{V}_{(N-r)}$ (annihilable by M) are called the *nil-spaces* of M. The matrix with non-trivial nil-spaces is called singular. Therefore, any matrix of the rank $1 \leq r < N$ is singular.

This scheme makes clear what the situation is when we want to solve (7.15) with respect to the x's, when the y's are considered as known. It is obvious that the consistency of (7.15) requires that

$$y^a \in V^r \quad , \qquad y_a \in V_r \ . \tag{7.21}$$

If these conditions are met, the solutions are determined with the accuracy

$$x^a \bmod v^a \in \bar{V}^{(N-r)} \quad , \qquad x_a \bmod v_a \in \bar{V}_{(N-r)} \ . \tag{7.22}$$

Or, more explicitly, if M^a_b has the form (7.13), and furthermore if y_a and y^a are given as

$$y^a = \sum_{i=1}^{r} \overset{i}{y} \, \underset{i}{\xi}^a \quad , \qquad y_a = \sum_{i=1}^{r} \underset{i}{y} \, \overset{i}{\eta}_a \ . \tag{7.23}$$

then the solutions of (7.15) are given as

$$x^a = \sum_{i=1}^{r} \underset{i}{\eta}^a \, \overset{i}{y} + \sum_{i=r+1}^{N} \underset{i}{\eta}^a \, \overset{i}{C} \quad ,$$

$$x_a = \sum_{i=1}^{r} \overset{i}{\xi}_a \, y_i + \sum_{i=r+1}^{N} \overset{i}{\xi}_a \, C_i \quad , \tag{7.24}$$

where $\overset{i}{C}$, C_i, $i = r+1, \ldots, N$ are arbitrary constants.

The form (7.13) of the matrix, as well as formula (7.24), which offers the solution of (7.15) when y's are of the form (7.23), has a theoretical value only. It gives a clear schematic insight into the meaning of matrices understood as transformations. But by no means may these formulae be considered an effective solution, when M^a_b is given by just numerical values of its independent components.

To some extent, however, the general scheme discussed above may be translated into the language of formulae free from the use of ξ's and η's. Indeed the statement $M_{[r]} \neq 0$ along with the form (7.13) implies that

$$0 = \underset{i_1}{\xi}^{[a_1} \ldots \underset{i_r}{\xi}^{a_r]} \overset{i_1}{\eta}_{b_1} \ldots \overset{i_r}{\eta}_{b_r} = M^{[a_1}_{b_1} \ldots M^{b_r]}_{a_r} \tag{7.25}$$

with the summation being over the i's from 1 to r. This, however, makes it obvious that the necessary and sufficient conditions for y^a and y_b to be the form (7.23) may be written as

$$M^{[a_1}_{b_1} \ldots M^{a_r}_{b_r} \, y^{a]} = 0 \quad , \quad M^{a_1}_{[b_1} \ldots M^{a_r}_{b_r} \, y_{b]} = 0 \, . \tag{7.26}$$

These are equivalent to

$$\mathfrak{M}^{[a]_{N-r}}{}_{[r].[b]_{N-r-1}\,[s]_1}\; y^s = 0 \;, \quad \mathfrak{M}^{[a]_{N-r-1}\,[s]_1}{}_{[r]\,[b]_{N-r}}\; y_s = 0 \;. \tag{7.27}$$

For $r = N - 1$ they become particularly simple:

$$\mathfrak{M}^a{}_{[N-1]s}\; y^s = 0 \;, \quad y_s\, \mathfrak{M}^s{}_{[N-1]b} = 0 \;, \quad \mathfrak{M}_{[N-1]} \in M_1 \;. \tag{7.28}$$

Now, although $M_{[r+1]} = 0$ implies $\underset{[r+1]}{M} = 0$, in general M $M_{[r]} \neq 0$ *does not* imply $\underset{[r]}{M} \neq 0$. Suppose, however, that we restrict ourselves to the special case in which

$$1 \leq r \leq N - 1 \;, \quad \underset{[r]}{M} \neq 0 \;. \tag{7.29}$$

[This is the case of M having the eigenvalue $M' = 0$ with the multiplicity exactly $N - r$; because $M\, D_r(M) = 0$ the algebraic defect of this multiplicity is $N - r - 1$; we discussed this in section V]. Under this assumption one can explicitly construct the solution of (7.15) by performing simple algebraic operations on the matrix M. Indeed, by contracting (7.27) one gets

$$\mathfrak{M}^a{}_{[r]s}\; y^s = 0 \;, \quad y_s\, \mathfrak{M}^s{}_{[r]b} = 0 \;, \quad \mathfrak{M}_{[r]} \in M_1 \;. \tag{7.30}$$

In general these are only necessary but not sufficient conditions of consistency [for $r = N-1$ they are also sufficient]. But under assumption (7.29) we shall demonstrate that they are sufficient.

Write (4.27) with $p = r,\ r+1$, remembering that $\underset{[r+1]}{\mathrm{M}} = 0$

$$\mathfrak{m}_{[r]} + \mathfrak{m}_{[r-1]} \cdot \mathrm{M} = \underset{[r]}{\mathrm{M}} \cdot \mathbf{1} \tag{7.31a}$$

$$\mathfrak{m}_{[r]} \cdot \mathrm{M} = 0 . \tag{7.31b}$$

Assumption (7.29) implies that $\mathfrak{m}_{[r]} \neq 0$, $\mathfrak{m}_{[r-1]} = 0$. If $\mathfrak{m}_{[r]} = 0$, (7.31a), together with $\underset{[r]}{\mathrm{M}} \neq 0$, would mean that M admits the inverse; this contradicts $r \leq N-1$. If $\underset{[r-1]}{\mathrm{M}} = 0$ we would have $\mathfrak{m}_{[r]} = \underset{[r]}{\mathrm{M}} \cdot \mathbf{1}$; using this in (7.31b) we get $\underset{[r]}{\mathrm{M}}\ \mathrm{M} = 0$ which contradicts $r > 1$ (when $\underset{[r]}{\mathrm{M}} \neq 0$).

Now, we claim that

$$y^a = \mathrm{M}^a_s\, x^s \qquad x^a = 1/\underset{[r]}{\mathrm{M}} \quad \mathfrak{m}^a_{[r-1]s}\ y^s + \mathfrak{m}^a_{[r]s}\ C^s \tag{7.32}$$

where C^s is an arbitrary vector. Indeed, letting M^b_a act on the x^a given by this formula shows that the contribution with C^s vanishes because of (7.31b). On the other hand, using (7.31a) and the assumption about y^a, (7.30), we have

$$\mathsf{M} \cdot 1/\underset{[r]}{\mathsf{M}} \cdot \mathsf{M}_{[r-1]} \cdot y = 1/\underset{[r]}{\mathsf{M}} \left(\underset{[r]}{\mathsf{M}} \cdot \mathbf{1} - \mathfrak{m}_{[r]} \right) \cdot y = y$$

(7.33)

where for simplicity we omitted indices. Therefore, (7.30) happens here to be the sufficient condition of consistency. In order to prove that it is also necessary, we do not need to contract the general (7.27); because of $y = \mathsf{M} \cdot x$ acting with the left-hand member of (7.31b) on x we see that $\mathfrak{m}_{[r]} \cdot y = 0$.

The term $\mathfrak{m}^{a}_{[r]s} C^{s}$ exhausts all the arbitrariness in x^{a}; one can prove this by showing that under assumption (7.29) $\mathfrak{m}_{[r]} \in \mathcal{M}_1$ is of the rank $N - r$, so that $\mathfrak{m}^{a}_{[r]s} C^{s}$ contains exactly $N - r$ linearly independent vectors. It will be simpler, however, to demonstrate this point later with the help of projection operators. Notice that (7.32) may also be written as

$$x = -\frac{D_{r-1}(\mathsf{M})}{D_r(0)} \cdot y + D_r(\mathsf{M}) \cdot C .$$

(7.34)

Of course, (7.15) (ii) has a parallel solution.

VIII. THE IDEM-POTENT MATRICES

The matrix M with the property that

$$\mathsf{M}^2 = \mathsf{M} \left(\to \mathsf{M}^n = \mathsf{M} \quad , \; n \geq 1 \right) \tag{8.1}$$

is said to be idem-potent. This condition is trivially satisfied by $\mathsf{M} = \alpha \mathbf{1}$, $\alpha = 0, 1$. The matrix $\mathsf{M} \neq 0, \mathbf{1}$ and which obeys (8.1) is said to be non-trivially idem-potent. For the non-trivially idem-potent M's, (8.1) serves as the minimal equation. Indeed, an equation of the first order is excluded; it could be only $\mathsf{M} = 0, \mathbf{1}$. Therefore $\bar{N} = 2$ and $\bar{D}_2(\lambda) = \lambda(\lambda - 1)$. This, however, implies that

$$D_N(\lambda) = \lambda^{N-r} (\lambda - 1)^r = (-)^N \sum_{k=0}^{r} (-\lambda)^{r-k} \binom{r}{k} \quad ,$$

$$1 \leq r \leq N - 1 \quad . \tag{8.2}$$

The number r -the multiplicity of the eigenvalue $\mathsf{M}' = 1$ (the other is $= 0$)- characterizes the structure of M. Indeed, one sees at once from (8.2) that the invariants $\mathsf{M}_{[p]}$ are

$$\underset{[p]}{M} = \binom{r}{p}, \quad p = 0, 1, \ldots, r; \quad \underset{[p]}{M} = 0 \text{ for } p > r. \tag{8.3}$$

The invariants $\overset{p}{M}$ according to $\overset{p}{M} = \sum_{i=1}^{N_0} n_i \, M_i^{\cdot p}$ are

$$\overset{0}{M} = N, \quad \overset{p}{M} = r \text{ for } p \geq 1. \tag{8.4}$$

(For $p = 0$ we always have $\overset{0}{M} = \sum n_i$). The invariants Δ_s are

$$\Delta_1 = N, \quad \Delta_2 = r(N - r); \quad \Delta_s = 0 \quad s > 2. \tag{8.5}$$

Let M be of the rank r'; as such it may be written as

$$M = \left\| \sum_{i=1}^{r'} \underset{i}{\xi}^{a} \; \underset{b}{\overset{i}{\eta}} \right\| \tag{8.6}$$

where the ξ's and η's are linearly independent. Using this form in (8.1) we find that

$$\sum_{i,j}^{r'} \underset{i}{\xi}^{a} \; \underset{b}{\overset{j}{\eta}} \left(\underset{s}{\overset{i}{\eta}} \; \underset{j}{\xi}^{s} - \delta_j^i \right) = 0. \tag{8.7}$$

Here using the linear independence we conclude that

$$\overset{i}{\underset{s}{\eta}} \; \underset{j}{\xi}^{s} = \delta^{i}_{j} \rightarrow \sum_{i=1}^{r'} \overset{i}{\underset{s}{\eta}} \; \underset{i}{\xi}^{s} = r' \; .$$

(8.8)

This implies that $Tr\left(M\right) = M^{s}_{s} = r'$. But $Tr\left(M\right) = \underset{[1]}{M} = \overset{1}{M} = r$

Therefore

$$r = r' \qquad (8.9)$$

and the multiplicity of the eigenvalue $M^{\prime} = 1$ coincides with the rank of the idem-potent matrix. Because of (8.8) the ξ's and η's may be identified with $\underset{i}{e}^{a}$ and $\overset{i}{e}_{a}$ which are a basis and its inverse. Spanning on $\underset{1}{e}^{a} \ldots \underset{r}{e}^{a}$ the sub-space $V^{(r)}$, on $\underset{r+1}{e}^{a} \ldots \underset{N}{e}^{a}$ the sub-space $V^{(N-r)}$ so that $V^{N} = V^{(r)} \oplus V^{(N-r)}$ we see that

$$x \in V^{(r)} \rightarrow x = M \, . \, x \; , \quad x \in V^{(N-r)} \rightarrow M \, . \, x = 0 \, .$$

(8.10)

Therefore the idem-potent M of the rank r may be understood as the transformation which projects any vector v into its part which belongs to some r-dimensional sub-space $V^{(r)}$. For this reason idem-potent matrices are also called projection operators. Of course, the parallel interpretation applies with respect to co-vectors.

Note that $M_\perp = \mathbf{1} - M$ is also idem-potent of the rank $N - r$. It projects vectors into the sub-space $V^{(N-r)}$. The following relations are true:

$$M_\perp M = 0 = M M_\perp , \quad M_\perp^2 = M_\perp , \quad M + M_\perp = \mathbf{1} . \tag{8.11}$$

IX. NIL-POTENT MATRICES

The study of the properties of nil-potent matrices constitutes an integral part of the theory of the canonical forms on the level of the standard matrices. The same is true on the level of our tensorial matrices where the standard theory of nil-potent matrices applies without any essential changes. For the sake of completeness we will sketch the theory in this section.

The matrix $\mathsf{M} \in M_1$ with the property that

$$\mathsf{M}^q = 0 \ , \ \mathsf{M}^{q-1} \neq 0 \ , \ q \geq 2 \tag{9.1}$$

is called *nil-potent*; the integer q is called the *index* of nil-potency.

We claim that $q \leq N$. Indeed, assume the opposite, i.e., $q > N \rightarrow q - 1 - N \geq 0$. But we always have $\sum_{p=0}^{N} \left(-\mathsf{M}\right)^{N-p} \underset{[p]}{\mathsf{M}} = 0$. Multiplying it by $\left(\mathsf{M}\right)^{q-1-0}$ and using (9.1) we get $\underset{[N]}{\mathsf{M}} = 0$. Therefore, the Hamilton-Cayley equation reduces to $\sum_{p=0}^{N-1} \left(-\mathsf{M}\right)^{N-p} \underset{[p]}{\mathsf{M}} = 0$. Multiplying it by $\left(-\mathsf{M}\right)^{q-1-1}$ we deduce that $\underset{[N-1]}{\mathsf{M}} = 0$. Repeating this process, we get in the l-the step, multiplying $D_N\left(\mathsf{M}\right) = 0$ by $\left(-\mathsf{M}\right)^{q-l-1}$ the conclusion that $\underset{[N-l]}{\mathsf{M}} = 0$. In the N-th step we would get (multiplying $D_N\left(\mathsf{M}\right) = 0$ by $\left(-\mathsf{M}\right)^{q-1-N}$ with the positive exponent) $\underset{[o]}{M} = 0$, which contradicts $\underset{[o]}{M} \underset{df}{=} 1$. Hence, $2 \leq q \leq N$.

Let $2 \leq q \leq N$. We claim that $M^q = 0$ is the minimal equation of M. If it were not, the following minimal equation of the form $\sum_{p=0}^{N} \left(- M\right)^{\bar{N}-p} \underset{[p]}{\bar{M}} = 0$ would be true, where $\underset{[o]}{M} = 1$, $q - \bar{N} > 0$ or equivalently $q - 1 - \bar{N} \geq 0$. Repeating exactly the same process as in the proof that $q \leq N$, we would arrive at the contradiction $\underset{[o]}{\bar{M}} = 1$. Hence $\bar{N} = q$. Because the minimal polynomial is unique, we therefore have

$$\bar{D}_q(\lambda) = \lambda^q \rightarrow D_N(\lambda) = \lambda^N . \tag{9.2}$$

Therefore, the nil-potent matrix has the single eigenvalue $M' = 0$ with the multiplicity $n = N$ and the algebraic defect $\delta n = N - q$. From (9.2) it follows immediately that

$$\underset{[p]}{M} = \delta_{p,o} \ , \ p = 0,1,\dots N \ , \ \overset{p}{M} = N\,\delta_{p,0} \ , \ p = 0,1,\dots,$$

$$\Delta_s = N\,\delta_{s,1} \ , \ s = 1,2,\dots \tag{9.3}$$

The inverse statement is also true: the matrix with vanishing $\underset{[1]}{M} \dots \underset{[N]}{M}$ is nil-potent. Indeed, such a matrix has $D_N(\lambda) = \lambda^N$. Hence, $\bar{D}(\lambda)$ must be of form $\bar{D}_q(\lambda) = \lambda^q$, $2 \le q \le N$, if M **is not identically zero.** From the very concept of the minimal polynomial it follows that $M^{q-1} \neq 0$.

A general comment about the rank of nil-potent matrices : because $M_{[N]} = 0$ all nil-potents are singular, $r \le N-1$. Now, because of (9.3), equation (4.28) reads

$$\mathfrak{m}_{[p]} = (-)^p \, M^p \; , \quad p = 0,1,\dots,N-1 \; , \quad \mathfrak{m}_{[p]} \in {}^c\mathfrak{M}_1 \; . \tag{9.4}$$

This implies that $\mathfrak{m}_{[q-1]} \neq 0$. Because $\mathfrak{m}_{[q-1]}$ contains linearly $M_{[q-1]}$ this quantity cannot identically vanish. Hence

$$N-1 \ge r \ge q-1 \; . \tag{9.5}$$

Thus, the possible ranks of nil-potents are bounded according to (9.5). This restriction for $q = N$ implies that $r = N-1$. Also, for $q = N-1$ there is a special situation. Namely, according to (9.5) and (9.1) we have that ${}_1\mathfrak{m}_{[N-1]} = 0$ implies that $M_{[N-1]} = 0$. Hence, here $r = N-2$. But for the lower q's there exists more freedom for the possible r's.

Now, without entering into details of the proof, we shall quote the important theorem about nil-potent matrices which one proves in the linear algebra and which also remains true for the nil-potent matrices in our sense (this theorem is valid also if the considered matrices have elements from a field of numbers which not necessarily is algebraically closed).

If the $N \times N$ matrix M is nil-potent with the index $2 \le q \le N$ then, there exists a number l and numbers $q_1, q_2, \dots q_l$ such that

$$(i) = q_1 \geq q_2 \geq \ldots \geq q_l \geq 1$$

$$(ii) \quad q_1 + q_2 + \ldots + q_l = N \tag{9.6}$$

and there exist vectors $\underset{1}{v}^a, \underset{2}{v}^a \ldots \underset{l}{v}^a \in V^N$ such that the vectors

$$(iii) \quad \underset{(p,k)}{e} \underset{df}{=} M^k \underset{p}{v} \qquad \begin{array}{l} p = 1, 2, \ldots \\ k = 0, 1, \ldots, q_p - 1 \end{array}$$

form a basis in V^N; moreover

$$(iiii) \quad \underset{(p,q_p)}{e} = M^{q_p} \underset{p}{v} \equiv 0 .$$

The numbers $q_1, \ldots q_l$ with the properties specified above form the sequence of the arithmetic invariants of the nil-potent matrix.

This theorem is the key to the understanding of the nature of the nil-potent matrices. Observe that M acts on the vectors of the basis defined in (9.6) according to

$$M \underset{(p,k)}{e} = \underset{(p,k+1)}{e} \; ; \; M \underset{(p,q_p - 1)}{e} = 0 . \tag{9.7}$$

The choice of the vectors $v_1, \dots v_l$ is not unique; the arithmetic invariants, however, are uniquely determined.

Let the inverse basis to that in (9.6) be defined as

$$\underset{(p,k)}{e^s} \; \overset{(p',k')}{e_s} = \delta_p^{p'} \cdot \delta_k^{k'} \rightarrow \sum_{p=1}^{l} \sum_{k=0}^{q_p - 1} \underset{(p,k)}{e^a} \; \overset{(p,k)}{e_b} = \delta_b^a \tag{9.8}$$

This together with (9.7) implies that

$$M = \left\| \sum_{p=1}^{l} \sum_{k=1}^{q_p - 1} \underset{(p,k)}{e^a} \; \overset{(p,k-1)}{e_b} \right\| \tag{9.9}$$

from which we may conclude that

$$\overset{(p,k)}{e} M = \overset{(p,k-1)}{e} \quad ; \quad \overset{(p,0)}{e} M = 0 \; . \tag{9.10}$$

One also easily sees that

$$M^s = \left\| \sum_{p=1}^{l} \sum_{k=s}^{q_p - 1} \underset{(p,k)}{e^a} \; \overset{(p,k-s)}{e_b} \right\| , \quad s = 0, 1, \dots, q \; . \tag{9.11}$$

[We understand the sum where s is $>$ than $q_p - 1$ to be empty]. Now, let $\Theta(n)$ be defined as* $n \geq 0 \rightarrow \Theta(n) = 1$, $n \leq -1 \rightarrow \Theta(n) = 0$. Such a function may be represented as

$$\Theta(n) = \sum_{m=0}^{\infty} \delta_{n,m} = \sum_{m=0}^{\infty} \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^{n+1}} z^m = \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^{n+1}} \frac{1}{1-z} .$$

(9.12)

From the form of (9.11) and the linear independency of $\underset{(p,k)}{e}$ it is obvious that

$$r_s \underset{df}{=} \text{rank}\left(M^s\right) = \sum_{p=1}^{1} \Theta(q_p - 1 - s) \cdot (q_p - s) \quad s = 0, 1, \ldots, q-1$$

(9.13)

also

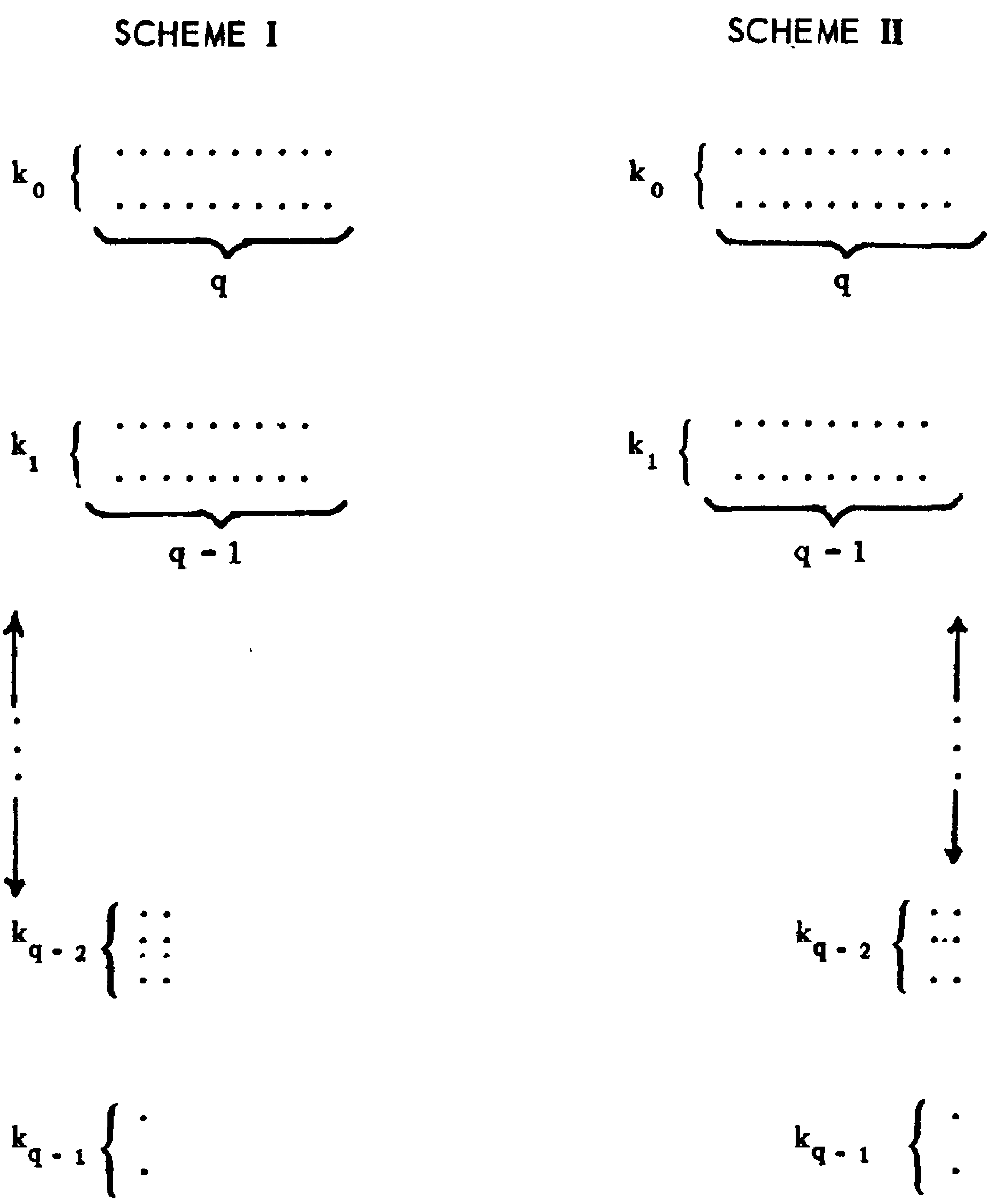
$$M^s \underset{(p,k)}{e} = \underset{(p,k+s)}{e} \;;\; \overset{(p,k)}{e} M^s = \overset{(p,k-s)}{e} .$$

(9.14)

The M^s therefore has nil-spaces of the dimension $n_s = N - r_s$. Formula (9.14), together with $\underset{(p,k)}{e} = 0$ for $k \geq q_p$ and $\overset{(p,k)}{e} = 0$ for $k \leq -1$, makes it obvious which vectors of the basis belong to these nil-spaces.

The numbers r_s , n_s are also arithmetic invariants of M . It is obvious that the sequences of q's and r's determine one another. Hence, it is only a matter of convention which of these is to be considered as fundamental.

The sequences of q's, r's, and n's, although important in themselves, are not the most convenient to handle. In order to obtain a simpler arithmetic description of the properties of M one may group the vectors of the basis $M^k v_p$ according to two schemes:



where the numbers k_s denote the number of the sequences of the type $M^k v_p$ which end with $M^{q-s} v_p$. Clearly, $k_0 \geq 1$ while the other k's are $k_s \geq 0$ equal (equal to zero in the case where corresponding sequences do not exist). Looking on our schemes vertically, in I the first line corresponds to v_p's, the second to $M v_p$, etc. In the second scheme the last line corresponds to $M^{q_p - 1} v_p$, the next (going from right to left) to $M^{q_p - 2} v_p$, etc. In both schemes M shifts a given line into its neighbor on the right; but it may happen that after shifting, a part of the given line will be annihilated (in the first scheme) or it will meet new participants (in the second scheme). The numbers $k_0 \geq 1$, $k_1, \dots, k_{q-1} \geq 0$ are of course arithmetic invariants of M. They must satisfy the obvious restriction

$$N = \sum_{i=0}^{q-1} (q - i)\, k_i \quad . \tag{9.15}$$

All other arithmetic invariants may be expressed through them. As far as the q's are concerned, it is obvious that any $k_s > 0$ gives rise to k_s numbers $q_i = q_s$ participating in the sequence $q = q_1 \geq q_2 \geq q_3 \geq \dots \geq q_l \geq 1$. It is also obvious that

$$l = \sum_{i=0}^{q-1} k_i \quad . \tag{9.16}$$

As far as the numbers r_s are concerned, it is obvious from (9.11) that only those $e_{(p,k)}$ participate in M^s which in the first scheme are located in columns starting from the s-th column (this being included). The number of vectors in the s-th column clearly is

$$z_s = \sum_{i=0}^{q-1-s} k_i \ , \quad s = 0,1,\dots,q-1 \ . \tag{9.17}$$

Therefore

$$r_s = \sum_{j=s}^{q-1} \sum_{i=0}^{q-1-j} k_i = \sum_{i=0}^{q-1-s} (q-s-i)\, k_i \ . \tag{9.18}$$

In particular $r_0 = \sum_{i=0}^{q-1} (q-i)\, k_i = N$, $r_{q-1} = k_0$. This implies that

$$n_s = N - r_s = s \sum_{i=0}^{q-1-s} k_i + \sum_{i=q-s}^{q-1} (q-1)\, k_i \tag{9.19}$$

which may easily be seen from scheme II, by remembering that its last $s+1$ columns form the nil-space of M^s.

The numbers k_s are particularly convenient in the problem of finding what is the number of the possible discrete structures of our nil-potent $N \times N$ matrix with the index q. It is obvious that any combination of $k_0 \geq 1$, $k_1, \dots, k_{q-1}$ which obeys (9.15) contributes to one possible discrete structure. Therefore, denoting the number of the possible discrete structures as $\begin{bmatrix} N \\ q \end{bmatrix}$ we have

$$\begin{bmatrix} N \\ q \end{bmatrix} = \sum_{k_0=1}^{\infty} \sum_{k_1=0}^{\infty} \cdots \sum_{k_{q-1}=0}^{\infty} \delta_{N, \sum_{i=0}^{q-1} (q-i) k_i}$$

(9.20)

But the same may be rewritten as

$$\begin{bmatrix} N \\ q \end{bmatrix} = \frac{1}{2\pi i} \oint_{C_0} \frac{1}{z^{N+1}} \sum_{k_0=1}^{\infty} \sum_{k_1=0}^{\infty} \cdots \sum_{k_{q-1}=0}^{\infty} (z^q)^{k_0} (z^{q-1})^{k_1} \cdots z^{k_{q-1}} .$$

(9.21)

For sufficiently small z (close to 0) the sums converge so that

$$\begin{bmatrix} N \\ q \end{bmatrix} = \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^{N-q+1}} \prod_{i=1}^{q} (1 - z^i)^{-1}$$

(9.22)

One can easily see that our $\begin{bmatrix} N \\ q \end{bmatrix}$ is the number of possible partitions of $N - q$ into numbers $\leq q$. Note that $\begin{bmatrix} N \\ q \end{bmatrix} = \begin{bmatrix} N \\ N-1 \end{bmatrix} = 1$ in agreement with the fact that $q = N$ implies $r = N - 1$, $q = N - 1$ implies $r = N - 2$. Note that for $q = 2$ (the extreme case for q) (9.22) yields

$$\begin{bmatrix} N \\ 2 \end{bmatrix} = m \text{ when } N = 2m \text{ or } 2m + 1 .$$

(9.23)

We shall conclude our remarks about nil-potent matrices with the statement that, although they do not admit the inverse, one may define to these the pseudo-inverse matrix which to some extent provides us with (not unique) inverse transformation. This matrix $M^{(-1)}$, we define by the conditions

$$M^{s} = M^{s}[M^{(-1)}]^{s} M^{s}, \quad [M^{(-1)}]^{s} M^{s} [M^{(-1)}]^{s} = M^{(-1)\ s}$$

9.24)

If the matrix M is non-singular these conditions are uniquely satisfied by $M^{(-1)} = M^{-1}$. In the case of the nil-potent M with the index q this chain of relations fails at $s = q - 1$.

From our previous analysis one can easily see that such $M^{(-1)}$ certainly exists. Indeed, the matrix:

$$M^{(-1)} = \left\| \sum_{p=1}^{l} \sum_{k=1}^{q_p - 1} \underset{(p,k-1)}{e^{a}} \overset{(p,k)}{e_b} \right\| \tag{9.25}$$

constructed from the basis defined in (9.6) and its inverse satisfies conditions (9.24) It is obvious that if $M^{(-1)}$ obeys (9.24) also $C M^{(-1)} C^{-1}$ will fulfill it if C is any non-singular matrix commuting with M.

The conditions (9.24) imply that the matrices

$$P^{s} = M^{s}[M^{(-1)}]^{s}, \quad Q^{s} = [M^{(-1)}]^{s} M^{s}$$

(9.26)

are idem-potent. They generate their idem-potent complements

$$P_\perp^s = \mathbf{1} - P^s \ , \quad Q_\perp^s = \mathbf{1} - Q^s \ . \tag{9.27}$$

In terms of these quantities, relations (9.24) say that

$$M^s \ Q_\perp^s = 0 \, , \quad P_\perp^s M^s = 0 \, , \quad Q_\perp^s [\, M^{(-1)} \,]^s = 0 \, , \, [\, M^{(-1)} \,]^s \ P_\perp^s = 0 . \tag{9.28}$$

Consider now the transformations of the type

$$y = M^s x \qquad s = 1,2,\ldots, q-1 \ . \tag{9.29}$$

The consistency of these equations for x requires

$$P_\perp^s \ y = 0 \ . \tag{9.30}$$

Therefore $y = P^s \ y$; now, decompose x as

$$x = \left(Q^s + Q_\perp^s \right) x = x_\parallel + x_\perp \quad . \tag{9.31}$$

Consequently, using $M^s \ Q_\perp^s = 0$, equation (9.29) reduces to

$$y = M^s x_{\parallel} . \tag{9.32}$$

Acting on it with $\left[M^{(-1)}\right]^s$ we have

$$\left[M^{(-1)}\right]^s y = Q^s x_{\parallel} = x_{\parallel} . \tag{9.33}$$

Therefore, the solution of (9.29) may be written as

$$x = \left[M^{(-1)}\right]^s y + Q^s_{\perp} v \tag{9.34}$$

where v is arbitrary. This formula justifies the term 'pseudo-inverse' as the name of $M^{(-1)}$. Of course, similar formulae are true when M is understood as a transformation in V_N. It is also obvious, because of the symmetry, that M may serve as the pseudo-inverse matrix to $M^{(-1)}$ when one has to solve $y = \left[M^{(-1)}\right]^s x$ with respect to x.

The last comment of this section: we may give to the numbers $r_s = \text{rank}\left(M^s\right)$, the independent arithmetic invariants of the nil-potent M, an alternative interpretation. Namely, there holds

$$M^s_{[k]} = M^{s\,a_1}_{[b_1} ::: M^{s\,a_k}_{b_k]} = \left[M_{[k]}\right]^s , \quad M_{[k]} \in \mathcal{M}_k . \tag{9.35}$$

($M^{s\,a}_{b}$ denotes s-th power of the matrix with indices a, b.)

It follows that the skew powers $M_{[r_s]} \in \mathcal{M}_{r_s}$ are nil-potent with the index $q(r_s) = s + 1$.

X. THE RESOLVENT AND THE MEROMORPHIC FUNCTIONS ON MATRICES.

In the end of section IV we saw that if $P(\lambda)$ is any polynomial of the letter λ with the property $P(\mathbf{M}) = 0$ it may serve for the explicit construction of the matrix $R(\lambda) = [\lambda - \mathbf{M}]^{-1}$ (for simplicity of notation we omit $\mathbf{1}$ at λ; we shall do so through all this section; this should not lead to any confusion). This matrix is called the resolvent. The most economic choice for $P(\lambda)$ in (4.42) is $P(\lambda) = \bar{D}_{\bar{N}}(\lambda)$; this choice gives us the representation of $R(\lambda)$ as a polynomial of the $\bar{N} - 1$ degree in the matrix $\mathbf{M}$

$$R(\lambda) \equiv \frac{1}{\lambda - \mathbf{M}} = \sum_{p=0}^{\bar{N}-1} \frac{\bar{D}_{\bar{N}-1-p}(\lambda)}{D_{\bar{N}}(\lambda)} \mathbf{M}^p . \tag{10.1}$$

The resolvent exists and is uniquely defined for any complex λ different from the eigenvalues of $\mathbf{M}$, $\mathbf{M}_i$'s. Therefore, the eigenvalues may be interpreted as the singularities of the resolvent. These singularities, zeros of $\bar{D}_{\bar{N}}(\lambda)$, form a set of isolated points on the complex plane of λ.

Note that construction (10.1), when convenient, may also be repeated on the basis of the characteristic polynomial $D_N(\lambda)$; on the right side of (10.1) we have, however, in this case, a polynomial of the $N - 1$ degree in $\mathbf{M}$.

Now, with the help of the concept of the resolvent we would like to define the concept of the meromorphic function on the matrix $\mathbf{M}$.

Let $F(z)$ be a meromorphic function of the complex z with possible singularities z_s, $s = 1,2,\ldots$ distributed arbitrarily on the complex plane which, --- is an essential assumption -- do not coincide with any of the singularities of the resolvent, i.e., with the eigenvalues M_i.

Now, consider the matrix, being a linear functional of $F(z)$ defined to the original matrix M, as

$$M[F(z)] = \frac{1}{2\pi i} \oint_C \frac{dz}{z - M} F(z) ; \tag{10.2a}$$

$$M[F(z)] = \sum_{p=0}^{\bar{N}-1} \left\{ \frac{1}{2\pi i} \oint_C dz \frac{\bar{D}_{\bar{N}-1-p}(z)}{\bar{D}_{\bar{N}}(z)} F(z) \right\} M^p , \tag{10.2b}$$

where the contour C is so chosen that it contains all M_i's (the singularities of the resolvent) but none of z_s (the singularities of $F(z)$). This matrix is well-defined and always exists.

We would like to show, by studying the consequences of the definition (10.2), that a so defined linear functional of $F(z)$ deserves the name of "the meromorphic function on the matrix M". This will be the main topic of this section, where we will use only the definition (10.2a) and the fact that the singularities of the resolvent are isolated.

In the next section we shall study the consequences of (10.2b), i.e., of the detailed structure of the resolvent, approaching that way the problem of the canonical forms of the matrix.

According to the definition of the contour C in (10.2a) it is obvious that (10.2a) may also be rewritten as

$$M[F(z)] = \sum_{i=1}^{N_0} M_i [F(z)] = \sum_{i=1}^{N_0} \frac{1}{2\pi i} \oint_{C_{M_i}} \frac{dz}{z - M} F(z) . \tag{10.3}$$

Where $C_{M_i^{\backprime}}$ are small contours around $M_i^{\backprime}$'s (which do not contain the singularities of $F(z)$); regions inside separate $C_{M_i^{\backprime}}$ are assumed to be disconnected [the C in (10.2a) may always be deformed into the sum of $C_{M_i^{\backprime}}$'s] • (10.3) defines uniquely the decomposition of $M_i\,[F(z)]$.

The "partial matrices" $M_i\,[F(z)]$ enjoy remarkable properties. Indeed, an identity holds:

$$\frac{1}{z_1 - M} \cdot \frac{1}{z_2 - M} = \frac{1}{z_2 - z_1}\left(\frac{1}{z_1 - M} - \frac{1}{z_2 - M}\right), \quad z_1 \neq z_2 , \tag{10.4}$$

which is a perfectly well-defined matrix equation with the proviso that z_1 and z_2 do not coincide with eigenvalues. Multiply it by $F(z_1)$ and $G(z_2)$ and integrate the output $\oint_{C_{M_i^{\backprime}}} dz_1 \oint_{C_{M_j^{\backprime}}} dz_2$, the contours being so close to eigenvalues that they do not contain singularities of $F(z_1)$ or $G(z_2)$. Because the domains inside of contours are disconnected and because of the structure of (10.4) we easily derive

$$i \neq j \rightarrow M_i\,[F(z)] \cdot M_j\,[G(z)] = 0 , \tag{10.5}$$

for any $F(z)$, $G(z)$. Now, let $C_{M_i^{\backprime}}$ and $\tilde{C}_{M_i^{\backprime}}$ be the contours around the same $M_i^{\backprime}$ so close to it that they do not contain singularities of $F(z)$, $G(z)$. Moreover, assume that $\tilde{C}_{M_i^{\backprime}}$ lies *outside* $C_{M_i^{\backprime}}$. Integrating (10.4) – multiplied by

$F(z_1)$, $G(z_2)$ - along these contours $\left(\oint_{C_{M_i^*}} dz_1 \quad \oint_{\tilde{C}_{M_i^*}} dz_2 \right)$ we derive

$$M_i\,[F(z)] \quad M_i\,[G(z)] =$$

$$= \left(\frac{1}{2\pi i}\right)^2 \left\{ \oint_{C_{M_i^*}} dz_1 \frac{F(z_1)}{z_1 - M} \oint_{\tilde{C}_{M_i^*}} dz_2 \frac{G(z_2)}{(z_2 - z_1)} + \right.$$

$$\left. + \oint_{C_{M_i^*}} dz_2 \frac{G(z_2)}{z_2 - M} \oint_{C_{M_i^*}} dz_1 \frac{F(z_1)}{(z_1 - z_2)} \right\} .$$

(10.6)

On the left side we simply wrote for the corresponding integrals the symbols $M_i\,[F(z)]$, because these clearly do not depend on the particular choice of contours. On the right side, in the first term, the integration over z_2 may be performed; because z_1 lies *inside* $C_{M_i^*}$ where $G(z_2)$ is analytic, the Cauchy formula may be applied here. In the second term we first execute the integration over z_1; but, because z_2 lies *outside* $C_{M_i^*}$ the integrand is holomorphic here so that the integral vanishes. Therefore, (10.6) reduces to

$$M_i\,[F(z)] \cdot M_i\,[G(z)] = M_i\,[F(z)\;G(z)] .$$

(10.7)

The formulae (10.5), (10.7) together imply the genera:

$$\mathsf{M}\,[F(z)] \cdot \mathsf{M}\,[G(z)] = \mathsf{M}\,[F(z)\;G(z)]\,, \tag{10.8}$$

which is valid for any $F(z)$, $G(z)$.

Now, we claim that no matter what M is, the following is always true

$$\mathsf{M}\,[1] = \frac{1}{2\pi i} \oint_C \frac{dz}{z - \mathsf{M}} = \mathbf{1}\,. \tag{10.9}$$

Indeed, let a be a complex number; we have

$$\left(a\mathsf{M}\right)[1] = \frac{1}{2\pi i} \oint_{C_a} \frac{dz}{z - a\mathsf{M}} = \frac{1}{2\pi i} \oint_C \frac{dz}{z - a\mathsf{M}} \underset{df}{=} \mathsf{M}\,(a) \tag{10.10}$$

where C_a contains the points $a\mathsf{M}_i^{\prime}$, which according to (5.13) are eigenvalues of the $a\mathsf{M}$. Of course, such a contour may be deformed into C. For any finite a the matrix $\mathsf{M}\,(a)$ certainly exists and is well-defined. In particular, obviously, $\mathsf{M}\,(0) = \mathbf{1}$. On the other hand for any $a \neq 0$ we also have

$$\left(a\mathsf{M}\right)[1] = \sum_{i=1}^{N_0} \frac{1}{2\pi i} \oint_{C_{a M_i^{\prime}}} \frac{dz}{z - a\mathsf{M}} = (z \to z\,a) =$$

$$= \sum_{i=1}^{N_0} \frac{1}{2\pi i} \oint_{C_{M_i}} \frac{dz}{z - \mathsf{M}} = \mathsf{M}\,[1] \; .$$

(10.11)

From these facts one concludes that $\lim_{a \to 0} (a\mathsf{M})\,[1] = \mathsf{M}\,[1]$. On the other hand, $\lim_{a \to 0} (a\mathsf{M})\,[1] = \mathsf{M}(0) = \mathbf{1}$. Hence $\mathsf{M}\,[1] = \mathbf{1}$ and (10.10) is true.

Now, the decomposition (10.3) together with (10.9) gives us a decomposition of the unit matrix :

$$\mathbf{1} = \sum_{i=1}^{N_0} \mathsf{E}_i \underset{df}{=} \sum_{i=1}^{N_0} \frac{1}{2\pi i} \oint_{C_{M_i}} \frac{dz}{z - \mathsf{M}} \quad .$$

(10.12)

The matrices $\mathsf{E}_i = \mathsf{M}_i\,[1]$ in consequence of (10.5), (10.7) fulfill a very simple algebra :

$$\mathsf{E}_i \cdot \mathsf{E}_j = \delta_{ij}\, \mathsf{E}_i \quad (i, j = 1, 2, \ldots, N_0 \text{ ; no summation over } i) ,$$

(10.13)

therefore they are idem-potent. Their structure and properties will be studied in the next section. For the purpose of this section the information contained in (10.12) and (10.13) is sufficient.

Consider now $\mathsf{M}\,[z]$; we have

$$\mathsf{M}[z] = \frac{1}{2\pi i} \oint_C \frac{z\,dz}{z - \mathsf{M}} = \frac{1}{2\pi i} \oint_C [\, z - \mathsf{M} + \mathsf{M}\,] \frac{1}{z - \mathsf{M}}\,dz =$$

$$= \frac{1}{2\pi i} \oint_C \mathbf{1}\,dz + \mathsf{M}\,\frac{1}{2\pi i} \oint_C \frac{dz}{z - \mathsf{M}} = 0 + \mathsf{M} \cdot \mathsf{M}[1] = \mathsf{M} \cdot \mathbf{1} = \mathsf{M}. \tag{10.14}$$

This fact, together with (10.8) implies that

$$\mathsf{M}[z^n] = \mathsf{M}^n, \quad n = 0, 1, \dots \quad . \tag{10.15}$$

This, however, means (because of the linearity of the functional $\mathsf{M}[F(z)]$) that if $P(z)$ is any polynomial

$$\mathsf{M}[P(z)] = P(\mathsf{M}) \,. \tag{10.16}$$

Now, take the obvious identity

$$\frac{1}{\lambda - z} \cdot \frac{1}{z - \mathsf{M}} - \frac{1}{\lambda - z} \cdot \frac{1}{\lambda - \mathsf{M}} = \frac{1}{\lambda - \mathsf{M}} \cdot \frac{1}{z - \mathsf{M}} \tag{10.17}$$

(essentially the same as (10.4)), and integrate it over C containing all M_i's but not the point $z = \lambda$. Doing so, we obtain

$$M \left[\frac{1}{\lambda - z}\right] = \frac{1}{\lambda - M} \quad . \tag{10.18}$$

This, together with (10.8) implies that

$$M \left[\frac{1}{(\lambda - z)^n}\right] = \left[\frac{1}{\lambda - M}\right]^n \quad , \tag{10.19}$$

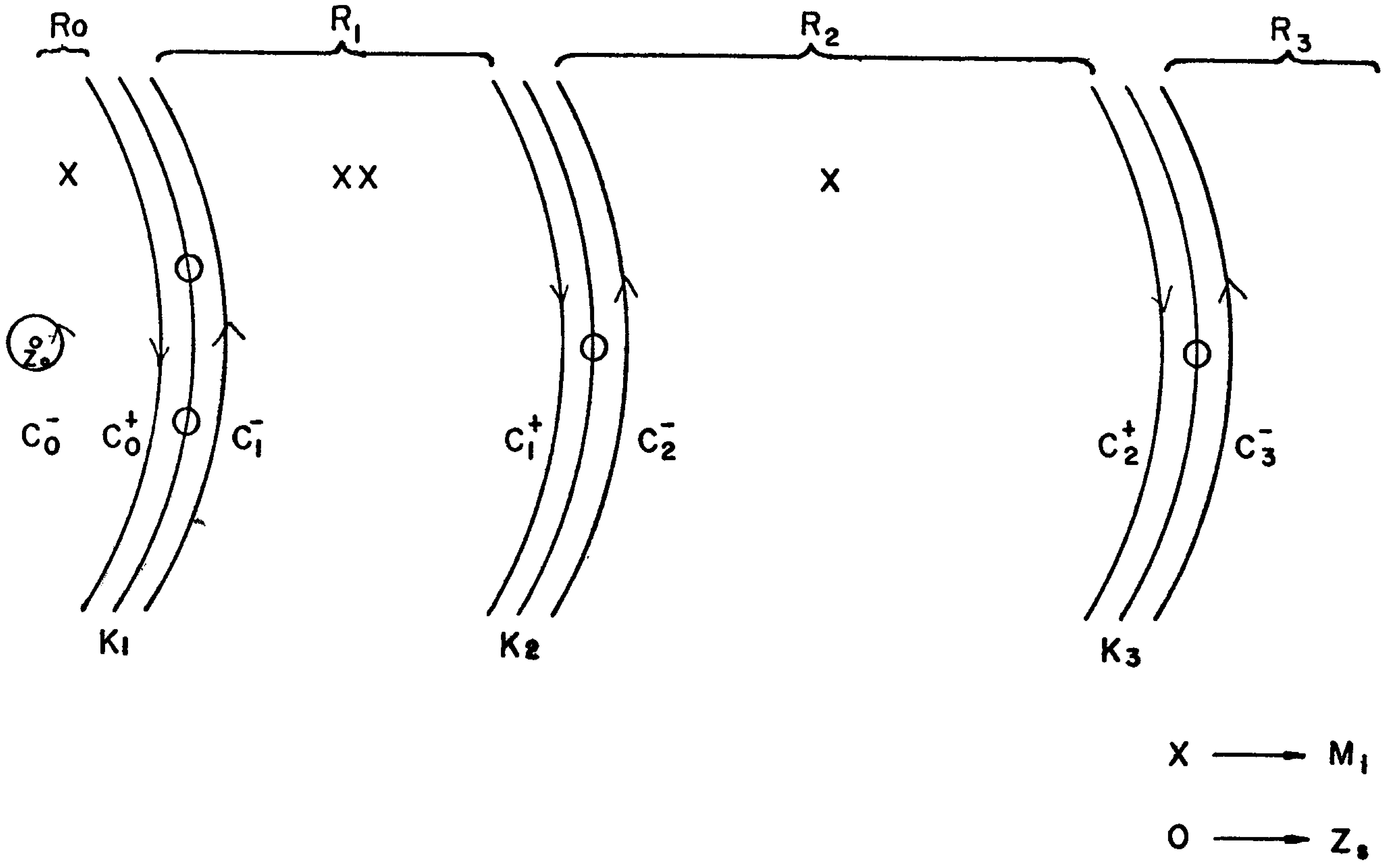
valid with the proviso $\lambda \neq M_i^{`}$, $i = 1, 2, \ldots N_0$. From this, one concludes that for any polynomial $P\left(\frac{1}{\lambda - z}\right)$ we have

$$M \left[P\left(\frac{1}{\lambda - z}\right)\right] = P\left(\frac{1}{\lambda - M}\right) \quad . \tag{10.20}$$

The facts contained in (10.16) and (10.20) give us a convenient starting point for approaching the Laurent-type development of our functionals.

Let z_0 be a point on the complex plane. It may be (but not necessarily is) one of the singular points of $F(z)$. Assume that it does not coincide with any of $M_i^{`}$'s. (Later we will see under what conditions this assumption is not needed.) Now, consider a family of circles $K_0, K_1, K_2, \ldots$ with the center at z_0. The K_0 coincides with the point z_0 itself (a degenerated circle). The K_1 has as one of its points the singularity of $F(z)$ which happens to be the closest to z_0. There may also be a few singularities of $F(z)$ on K_1 if they happen to be equidistant from z_0. The K_2 goes through the singularity of $F(z)$ (or a few singularities),

FIGURE 2



which is the next closest to z_0 etc. Now, consider for any circle K_i two circles C^+_{i-1} and C^-_i infinitesimally close to it, from the inside and the outside respectively. Between the circles C^-_i and C^+_i we have the domain of the complex plane, a ring, denoted as R_i. We shall assume that *all* singularities of the resolvent, M_i's, are somehow distributed *inside* of the family of rings defined above (none of M_i's lies on any of K_i's). By this assumption we simply intend tc remove some inconvenient *limiting locations* of z_0 from our *considerations. The situation is illus*trated by the map.

Now, because, as follows from our construction, the $F(z)$ is analytic in any of rings R_i, it may be represented there through the corresponding Laurent series

$$z \in R_i \rightarrow F(z) = \sum_{n=-\infty}^{+\infty} a^i_n (z - z_0)^n ;$$

$$a^i_n \underset{df}{=} \frac{1}{2\pi i} \oint_{C_i^{n/m_1}-} \frac{dz\, F(z)}{(z - z_0)^{n+1}}$$

(10.21)

of course, in general, different in each of the rings.

Let $\tilde{R}_1, \tilde{R}_2, \ldots \tilde{R}_m$ be those rings which are not empty, i.e., which contain at least one of the M_i's. Let $\tilde{E}_1, \tilde{E}_2, \ldots \tilde{E}_s$ be sums of E_i's (defined by (10.12)) which correspond to eigenvalues present in corresponding rings. Obviously

$$\sum_{k=1}^{m} \widetilde{E}_k = \mathbf{1} \quad , \quad \widetilde{E}_k \cdot \widetilde{E}_l = \delta_{kl} \widetilde{E}_k \quad .$$

(10.22)

After these preparations, we may approach the problem of the quasi-Laurent series for $\mathsf{M}[F(z)]$. First of all, we claim that by appropriately deforming the contour C containing $\grave{\mathsf{M}}_i$'s but none of z_s we may write:

$$\mathsf{M}[F(z)] = \frac{1}{2\pi i} \oint_C \frac{dz\, F(z)}{z - \mathsf{M}} = \sum_{k=1}^{m} \frac{1}{2\pi i} \oint_{\widetilde{C}_k} \frac{dz\, F(z)}{z - \mathsf{M}}$$

(10.23)

where C_k consists of two circles C_k^- and C_k^+ which form the frontiers of the 'non-empty' ring $\widetilde{R}_k$. (The ring C_k^- has to be traversed in a clockwise direction.) These circles, however, still may be moved a little bit inside of the ring where the uniformly convergent development (10.21) of the $F(z)$ is already valid.

[The shifted rings which still contain between them all the $\grave{\mathsf{M}}_i$'s from $\widetilde{R}_k$ we denote as $\widetilde{C}_k'$]. Therefore, if by $\tilde{a}_n^k$ we understand the coefficients of the Laurent development of $F(z)$ in $\widetilde{R}_k$, we may write

$$\mathsf{M}[F(z)] = \sum_{k=1}^{m} \frac{1}{2\pi i} \oint_{\widetilde{C}_k'} \frac{dz}{z - \mathsf{M}} \sum_{n=-\infty}^{+\infty} \tilde{a}_n^k (z - z_0)^n \quad .$$

(10.24)

We may here, however, sensibly interchange the order of the integration and summation. The Laurent series uniformly converges, so that the same is true for each essentially different component of the matrix considered. Hence

$$\mathsf{M}\,[F(z)] = \sum_{k=1}^{m} \sum_{n=-\infty}^{+\infty} \tilde{a}_n^k \frac{1}{2\pi i} \oint_{\tilde{C}'_k} \frac{dz\,(z - z_0)^n}{z - \mathsf{M}} \quad . \tag{10.25}$$

Now, comes a simple trick: because of the properties of our idem-potent $\tilde{E}_k$, clearly $\tilde{E}_k \oint_C \frac{dz\; G(z)}{z - \mathsf{M}} = \oint_{\tilde{C}'_k} \frac{dz\; G(z)}{z - \mathsf{M}}$.

Therefore, with the help of these matrices we rewrite (10.25)

$$\mathsf{M}\,[F(z)] = \sum_{k=1}^{m} \sum_{n=-\infty}^{+\infty} \tilde{a}_n^k\, \tilde{E}_k\, \frac{1}{2\pi i} \oint_{C} \frac{dz\,(z - z_0)^n}{z - \mathsf{M}} \quad . \tag{10.26}$$

In the last step of our construction we apply (10.16), (10.20) for the negative and the positive powers, obtaining

$$\mathsf{M}\,[F(z)] = \sum_{k=1}^{m} \sum_{n=-\infty}^{+\infty} \tilde{a}_n^k\, \tilde{E}_k \cdot (\mathsf{M} - z_0)^n \quad . \tag{10.27}$$

where the negative powers are to be understood as powers of the resolvent.

This result represents the quasi-Laurent development of $\mathsf{M}\,[F(z)]$ which generalizes (10.16), (10.20) for the case of arbitrary meromorphic functions. As it follows from its derivation, the series (10.27) is convergent in the sense that series for each essentially different element of the matrix $\mathsf{M}\,[F(z)]$ are convergent. Our quasi-Laurent development of $\mathsf{M}\,[F(z)]$ "around z_0" (i.e., "around

1 z_0") has the peculiarity that, in general, the coefficients of Laurent series for $F(z)$ from a few different rings are present in it. This is due to the fact that the matrix is not just one number but *many*. The corresponding Laurent series has to participate in (10.27) when there are present some eigenvalues in rings where the Laurent development holds. The matrix is, so to say, 'partially' present in these rings. The *idem-potent* 'projection operator' $\tilde{\mathsf{E}}_k$ extracts from the corresponding matrices exactly that part which concerns the eigenvalues present in the given ring.

Formula (10.27) contains a few interesting special cases. When *all* eigenvalues are located in *one* ring it obviously reduces to

$$\mathsf{M}\,[F(z)] = \sum_{n=-\infty}^{+\infty} a_n \left(\mathsf{M} - z_0\right)^n \tag{10.28}$$

where a_n's are Laurent coefficients from the ring in question. This is more like the familiar Laurent series. In particular, it may happen that the principal part of this development vanishes ($a_n = 0$ for $n < 0$). The $F(z)$ is here holomorphic in a circle around z_0 which contains all the eigenvalues.

In this case (10.28) may be rewritten as

$$\mathsf{M}\,[F(z)] = \sum_{n=0}^{\infty} \frac{1}{n!} F^{(n)}(z_0) \left(\mathsf{M} - z_0\right)^n \quad . \tag{10.29}$$

Now, summing up, we see that formula (10.27), which has as special cases (10.28), (10.29), (10.16), (10.20), entirely justifies another interpretation of our functional $\mathsf{M}\,[F(z)]$. We may understand it generally as the meromorphic function *on the matrix*, $F\left(\mathsf{M}\right)$; the consistency of such an interpretation is assured by (10.8).

For these reasons we shall also write

$$F\left(\mathsf{M}\right) \underset{df}{=} \mathsf{M}\,[F(z)] \ . \tag{10.30}$$

In all this section we kept M fixed but $F(z)$ variable. Of course, one can proceed otherwise, keeping $F(z)$ fixed but M variable with the proviso that its eigenvalues should not coincide with singularities of $F(z)$. The study of such relations lies outside the scope of this paper.

XI. THE CANONICAL FORMS, EIGENVECTORS

In the previous section we studied the consequences of our definition of the meromorphic function on the matrix without using the detailed structure of the resolvent. In this section we shall approach our $F(M) = M[F(z)]$ from the other angle, exploring the consequences of (10.26).

First rewrite (10.26) in the form where C is deformed into the sum of contours $C_{M_i^`}$, each around a separate singularity, surrounding it so close that inside of it, $F(z)$ is holomorphic; of course, regions inside $C_{M_i^`}$, $C_{M_j^`}$ are disconnected if $i = j$.

$$F(M) = \sum_{p=0}^{\tilde{N}-1} \left\{ \sum_{i=1}^{N_0} \frac{1}{2\pi i} \oint_{C_{M_i^`}} dz\ F(z)\ \frac{\bar{D}_{\tilde{N}-1-p}(z)}{\bar{D}_{\tilde{N}}(z)} \right\} M^p \ .$$

(11.1)

Because inside each $C_{M_i^`}$ the $F(z)$ is analytic, it there has the uniformly convergent Taylor's series

$$F(z) = \sum_{n=0}^{\infty} \frac{1}{n!} F^{(n)}\left(M_i^`\right)\left(z - M_i^`\right)^n, \ i = 1,2,\ldots,N_0 \ .$$

(11.2)

Substitute it into (11.1); the interchange of order of the integration and summation is legitimate. In effect we get

$$F\left(M\right) = \sum_{i=1}^{N_0} \sum_{n=0}^{\infty} \frac{1}{n!} F^{(n)}\left(M_i^{`}\right) E_i^{(n)} \tag{11.3}$$

where the matrices $E_i^{(n)}$ are defined as

$$E_i^{(n)} \underset{df}{=} \frac{1}{2\pi i} \oint_{C_{M_i^{`}}} dz \frac{\left(z - M_i^{`}\right)^n}{\bar{D}_{\bar{N}}(z)} \cdot \sum_{p=0}^{\bar{N}-1} \bar{D}_{\bar{N}-1-p}(z) \cdot M^p . \tag{11.4}$$

Define now the *polynomials* of the complex λ:

$$E_i^{(n)}(\lambda) \underset{df}{=} \frac{1}{2\pi i} \oint_{C_{M_i^{`}}} dz \frac{\left(z - M_i^{`}\right)^n}{\bar{D}_{\bar{N}}(z)} \sum_{p=0}^{\bar{N}-1} \bar{D}_{\bar{N}-1-p}(z) \lambda^p \tag{11.5}$$

of course, $E_i^{(n)}\left(M\right) = E_i^{(n)}$. But these polynomials may easily be reduced to a more plausible form. Using (4.40) we may rewrite them as

$$E_i^{(n)}(\lambda) = \frac{1}{2\pi i} \oint_{C_{M_i^{`}}} dz \frac{\left(z - M_i^{`}\right)^n}{z - \lambda} \left(1 - \frac{\bar{D}_{\bar{N}}(\lambda)}{\bar{D}_{\bar{N}}(z)}\right) . \tag{11.6}$$

If we restrict ourselves to λ *outside* $C_{M'_i}$, the first term does not contribute to the integral. In the second, we take $\bar{D}_{\bar{N}}(z)$ in the factorized form (5.7). Therefore, in the notation

$$\bar{D}^i(z) = \prod_{j=1}^{N_0} \left(z - M'_j\right)^{q_j}, \; j \neq i \; . \tag{11.7}$$

[The minimal polynomial with extracted factor $\left(z - M'_i\right)^{q_i}$; with the suffix i up, to distinguish it from the incomplete minimal polynomials $\bar{D}_k(z)$] the formula (11.6) reduces to

$$E_i^{(n)}(\lambda) = \frac{1}{2\pi i} \oint_{C_{M'_i}} \frac{dz}{\left(z - M'_i\right)^{q_i - n}} \cdot \frac{1}{\lambda - z} \cdot \frac{\bar{D}_{\bar{N}}(\lambda)}{\bar{D}^i(z)} =$$

$$= \frac{1}{(q_i - 1 - n)!} \left(\frac{d}{dz}\right)^{q_i - 1 - n} \frac{1}{\lambda - z} \; \frac{1}{\bar{D}^i(z)} \Bigg|_{z = M'_i} \bar{D}_{\bar{N}}(\lambda)$$

$$= \left(\lambda - M'_i\right)^n \sum_{k=0}^{q_i - 1 - n} \bar{A}_i^{(k)} \left(\lambda - M'_i\right)^k \bar{D}^i(\lambda) \; ;$$

$$\bar{A}_i^{(k)} = \frac{1}{k!} \left(\frac{d}{dz}\right)^k \frac{1}{\bar{D}^i(z)} \Bigg|_{z = M'_i} \; . \tag{11.8}$$

The second and third lines make sense only for $q_i - 1 - n > 0$. The first line clearly shows that for $n > q_i$ the $E_i^{(n)}(\lambda)$ do vanish.

Taking this fact into account, we rewrite (11.3) as

$$F(M) = \sum_{i=1}^{N_0} \sum_{n=0}^{q_i - 1} \frac{1}{n!} F^{(n)} M_i^{`} E_i^{(n)} \tag{11.9}$$

so that only finite derivatives of $F(z)$ may participate in this formula. But all the $E_i^{(n)}$ which appear in this formula are already non-trivial as may be seen from their explicit form (11.8) [after substituting $\lambda \to M$] which demonstrates that they are polynomials of the $(\bar{N} - 1)$th degree in M with non-zero coefficient at $M^{\bar{N}-1}$.

Specializing (11.9) for $F(z) \equiv 1$ and comparing (10.13) with (19.12), we conclude that the matrices E_i introduced before coincide with $E_i^{(0)}$:

$$E_i = E_i^{(0)} . \tag{11.10}$$

The explicit form of $E_i^{(n)}(\lambda)$, along with the definition of $\bar{D}^i(\lambda)$, manifestly shows that the eigenvalues of the matrix $E_i^{(n)}$, which according to (5.13) are just $E_i^{(n)}\left(M_j^{`}\right)$, $j = 1, 2, \ldots, N_0$ are given as

$$E_i^{(n)}\left(M_j^{`}\right) = \delta_{ij}\ \delta_{n0}\ , \quad i,j = 1,2,\ldots,N_0\ ;$$

$$n = 0,1,\ldots,q_i - 1\ . \tag{11.11}$$

This has important consequences. First, according to (10.15):

$$T_r\left(E_i\right) = T_r\left(E_i^{(0)}\right) = \sum_{j=1}^{N_0} n_j\ E_i^{(0)}\left(M_j^{`}\right) = n_i\ . \tag{11.12}$$

But the trace of an idem-potent matrix is equal to its rank. Hence:

$$\text{rank}\ \left(E_i\right) = n_i = \text{multiplicity of}\ M_i^{`}\ . \tag{11.13}$$

Secondly, because for $n > 1$ all of the eigenvalues of $E_i^{(n)}$ vanish, therefore the invariants of these matrices also vanish. Consequently, all matrices $E_i^{(n)}$ for $n \gtrsim 1$ are nil-potent.

Now, because $E_i^{(n)}(\lambda)$ is proportional to $\left(\lambda - M_i^{`}\right)^n\ \bar{D}^i(\lambda)$ therefore $\left(\lambda - M_i^{`}\right)^{q_i - n}\ E_i^{(n)}(\lambda)$ certainly contain $\bar{D}_{\bar{N}}(\lambda)$ as a factor. Consequently

$$\left(M - M_i^{`} \right)^{q_i - n} E_i^{(n)} = 0 \, , \quad n = 0,1,\ldots,q_i - 1 \, .$$

(11.14)

With the help of this formula and the third line of (11.8) one easily finds

$$E_i^{(n)} = \left(M - M_i^{`} \right)^n E_i \quad .$$

(11.15)

This form plus the algebra of E_i's at once implies that the $E_i^{(n)}$'s obey the simple algebra:

$$E_i^{(n)} \cdot E_j^{(m)} = \delta_{ij} \, E_i^{(n+m)} \, , \quad n \geq q_i \rightarrow E_i^{(n)} = 0 \; .$$

(11.16)

Also:

$$[E_i^{(n)}]^m = E_i^{(nm)} = 0 \, , \quad \text{for} \;\; nm \geq q_i \; .$$

(11.17)

which forms a more explicit expression of the fact that $E_i^{(n)}$ for $n \geq 1$, if non-trivial, are nil-potent.

Finally, we may define a set of nil-potent matrices:

$$M_i = (M - \mathbf{1}\, M_i^{`})\, E_i \;;$$

$$M_i^{q_i} = 0 \;,\; M_i^{q_i - 1} \neq 0 \;,\; 1 = 1, 2, \ldots, N_0 \quad .$$

(11.18)

The formula (11.9) may now be rewritten as

$$F(M) = \sum_{i=1}^{N_0} F_i(M) \quad ;$$

$$F_i(M) = \sum_{n=0}^{q_i - 1} \frac{1}{nI}\, F^{(n)}(M_i^{`})\; M_i^n\; E_i \quad .$$

(11.19)

This may be understood as the spectral decomposition of $F(M)$. The different eigenvalues of M, i.e., $M_i^{`}$, provide us with the 'rough' structure of $F(M)$ (its decomposition into $F_i(M)$). Simultaneously, if one looks upon $F(M)$ as a transformation in V^N, this 'rough' structure corresponds to the decomposition of V^N into a sum of n_i-dimensional sub-spaces V^{n_i} where F_i M operate non-trivially:

$$(v \in V^{n_i}) \to (v = E_i v) \;;\; V^N = V^{n_1} \oplus V^{n_2} \oplus \ldots ; \oplus V^{n_{N_0}} \quad .$$

(11.20)

Now, each $F_i(M)$ has its own 'subtle' structure determined by the arithmetic invariants of the nil-potent matrix M_i. In a sense, M_i may be understood as a matrix of the order $n_i \times n_i$, because $i \neq j$ implies $M_i V^{n_j} = 0$.

Gathering together that which was established about the nil-potent matrices in section IX, we may claim that there exists a number l_i and numbers

$$q_i = {}^1q_i \geq {}^2q_i \geq \dots \geq {}^{l_i}q_i \geq 1 \tag{11.21a}$$

such that

$$n_i = {}^1q_i + {}^2q_i + \dots {}^{l_i}q_i \tag{11.21b}$$

and there exist vectors ${}^1v_i$, ${}^2v_i$, ..., ${}^{l_i}v_i$ such that the vectors

$$\underset{(i,p,k)}{e} = M_i^k \; {}^pv_i \qquad \begin{array}{l} p = 1,2,\dots, l \\ k = 0,1,\dots, {}^pq_i - 1 \end{array} \tag{11.21c}$$

form the basis of V^{n_i}; moreover:

$$M_i \underset{(i,p,{}^pq_i-1)}{e} = M_i^{{}^pq_i} \cdot {}^pv_i = 0 . \tag{11.21d}$$

This basis has the property that

$$M_i \underset{(i,p,k)}{e} = \underset{(i,p,k+1)}{e} \quad \text{and where} \quad \underset{(i,p,{}^{p}q_i)}{e} = 0 . \tag{11.22}$$

The numbers ${}^{p}q_i$ with the properties (11.21a-b) are the arithmetic invariants which describe the subtle structure of the matrix M.

The statement contained in (11.21) is the literal repetition of the theorem (9.6) but as applied to the nil-potent M_i which operates non-trivially only on V^{n_i}, hence the necessity of adding the suffix i to all quantities concerned.

Now, similarly as was done in section IX we introduce the inverse basis to $\underset{(i,p,k)}{e}$ by

$$\underset{(i,p,k)}{e^s} \overset{(i',p',k')}{e} = \delta^{i'}_{i} \delta^{p'}_{p} \delta^{k'}_{k} \longrightarrow$$

$$\mathbf{1} = \left\| \sum_{i=1}^{N_0} \sum_{p=1}^{l_i} \sum_{k=0}^{{}^{p}q_i - 1} \underset{(i,p,k)}{e^a} \overset{(i,p,k)}{e_b} \right\| \tag{11.23}$$

Generalizing the considerations of section IX we easily find that the quantities $E^{(n)}_i$ may be represented through the vectors of our basis and its inverse as

$$E_i^{(n)} = \sum_{p=1}^{l_i} \sum_{k=n}^{p_{q_i-1}} e^a_{(i,p,k)} \; {}^{(i,p,k-n)}e_b \qquad \begin{array}{l} i = 1,2,\ldots,N_0 \\ \\ n = 0,1,\ldots,q_i - 1 \end{array}$$

(11.24)

Of course, all sums over k where $n > p_{q_i-1}$ are to be understood as empty. Using this representation of $E_i^{(n)}$ in (11.9) we may write $F\ M$ in the form

$$F(M) = \left\| \sum_{i=1}^{N_0} \sum_{n=0}^{q_i-1} \frac{1}{n!} F^{(n)}\left(M_i^{`}\right) \sum_{p=1}^{l_i} \sum_{k=n}^{p_{q_i-1}} e^a_{(i,p,k)} \; {}^{(i,p,k-n)}e_b \right\| .$$

(11.25)

This form of $F(M)$ may be called the Jordan's canonical form of the meromorphic $F(z)$ on M; it may be interpreted in the standard way. Namely, the quantities ${}^{(i,p,k)}e_a \; F(M)^a_b \; e^b_{(i',p',k')}$ which entirely describe the matrix $F(M)$ may be ordered into a $N \times N$ table, the columns and rows of which correspond to (i,p,k), (i',p',k'), properly ordered. This table clearly splits into a sequence of non-trivial $n_i \times n_i$ squares along the diagonal which correspond to the different eigenvalues $M_i^{`}$ (Jordan's cells). A given $n_i \times n_i$ square below the diagonal has only zeros. On the diagonal in all $n_i$ places is $F\left(M_i^{`}\right)$. In the first line above the diagonal (and parallel to it) there is first the sequence of $q_i - 1 = {}^1q_i - 1$ of $\frac{1}{1!} F^{(1)}\left(M_i^{`}\right)$'s, followed by one zero (starting from above). Then the sequence of ${}^2q_i - 1$ of $\frac{1}{1!} F^{(1)}\left(M_i^{`}\right)$'s followed by zero, and so on. In the next line above, parallel to the diagonal, there are located sequences of $\frac{1}{2!} F^{(2)}\left(M_i^{`}\right)$'s. The first $q_i - 2 \equiv {}^1q_i - 2$ of these followed by two zeros, then ${}^2q_i - 2$ followed by two zeros, etc. As may easily be generalized from the first three lines discussed, we meet a similar situation in higher lines above the diagonal (and parallel to it). E.g., the typical 'Jordan's cell' with $n_i = 6$, ${}^1q_i = 3 = {}^2q_i$ is visualized in figure I.

FIGURE II.

$$\begin{matrix}
F\left(M^{`}\right) & F^{(1)}\left(M^{`}\right) & \frac{1}{2}F^{(2)}\left(M^{`}\right) & 0 & 0 & 0 \\
0 & F\left(M^{`}\right) & F^{(1)}\left(M^{`}\right) & 0 & 0 & 0 \\
0 & 0 & F\left(M^{`}\right) & 0 & 0 & 0 \\
 & & & & & \\
0 & 0 & 0 & F\left(M^{`}\right) & F^{(1)}\left(M^{`}\right) & \frac{1}{2}F^{(2)}\left(M^{`}\right) \\
0 & 0 & 0 & 0 & F\left(M^{`}\right) & F^{(1)}\left(M^{`}\right) \\
0 & 0 & 0 & 0 & 0 & F\left(M^{`}\right)
\end{matrix}$$

For $F(z) = z$, i.e., $F\left(M\right) = M$ our scheme reduces strictly to the Jordanian picture, with eigenvalues located along the diagonal and descending sequences of 1's followed by zero above the diagonal line, all other places in the table are occupied by zeros.

Form (11.25) of the $F(M)$ is particularly convenient for studying the *eigenvectors* of the matrix $F(M)$, i.e., vectors $v_\lambda \neq 0$ such that

$$\left(F(M) - \lambda \mathbf{1} \right) \cdot v_2 = 0 .$$

(11.26)

It is obvious that the only admissible λ for which such a non-trivial v_λ may exist form the sequence: $\lambda = F(M_i^`)$, $i = 1,2,\ldots, N_0$. (In general, some of these may coincide). Now, the structure of (11.25) plus the linear independence of the vectors of our basis make it obvious that to the eigenvalue $\lambda = F(M_i^`)$, if $F^{(1)}(M_i^`) \neq 0$, belong *only* the vectors $e_{(i,p,{}^p q_i - 1)}$ as linearly independent eigenvectors, i.e., exactly $l_i$ of them. More generally, one easily sees that if $F^{(1)}(M_i), \ldots, F^{(s-1)}(M_i^`)$ all vanish but beginning with s: $F^{(s)}(M_i^`) \neq 0$ $(1 \leq s \leq q_i - 1)$ then $e_{(i,p,{}^p q_i - 1)}$, $e_{(i,p,{}^p q_i - 2)}$, ..., $e_{(i,p,{}^p q_i - s)}$ are the eigenvectors of $F(M)$ corresponding to $\lambda = F(M_i^`)$. When counting the number of these vectors one must take into account that for $k < 0$ the $e_{(i,p,k)}$ do not exist. Hence, the total number of these vectors is:

$$\sum_{m=0}^{s=1} \sum_{p=1}^{l_i} \Theta({}^p q_i - 1 - m) = n_i - \sum_{p=1}^{l_i} ({}^p q_i - s)\, \Theta({}^p q_i - 1 - s) .$$

(11.27)

For $s = q_i$ (i.e., when all $F^{(1)}(M_i^`), \ldots, F^{(q_i - 1)}(M_i^`)$ which participate in the given cell vanish) this number is just n_i, as was to be expected. Note that comparing (11.27) with (9.13) we may conclude that the total number of eigenvectors considered is invariantly given as

$$n_i - \text{rank}\left\{ M_i^s \right\} \quad . \tag{11.28}$$

Although the basis used in (11.25) is not unique, the numbers ${}^p q_i$ or (11.27), (11.28) are arithmetic invariants independent on the particular choice of the basis. The number (11.28) with the simple interpretation of the number of eigenvectors belonging to $\lambda = F(M_i^{`})$ when $F^{(k)}(M_i^{`}) = 0$ for $k = 1,2,\ldots, s-1$ but $F^{(s)}(M_i^{`}) \neq 0$ gives us an especially plausible interpretation of the arithmetic invariants which describe the 'subtle' structure of the matrix. Of course, the numbers (11.28) may be understood as the fundamental arithmetic invariants; all others may be expressed through them, as follows from the considerations made in section IX.

Here is the place to explain why we carried out the 'spectral' analysis on the level of $F(M)$ instead doing so on the level of M itself. Our $F(M)$ matrices generally contain $\bar{N}$ continuous parameters $[F^{(n)}(M_i^{`}), i = 1,2,\ldots, N_0;$ $n = 0,1,\ldots, p_i - 1]$ and by dealing with them we are able to give all the vectors of the basis appearing in the canonical form (11.25) the interpretation of the eigenvectors in the simple sense of (11.26). By doing so we avoid the necessity of considering the so-called principal vectors (or multiple eigenvectors) defined by $M_i^n v = 0$, where n is an integer, which are somewhat artificial creatures. Of course, the multiple eigenvectors are implicitly hidden in our formalism. But, in our discussion of the canonical forms, nothing forces us to introduce them explicitly.

XII. SOME SIMPLE FUNCTIONS ON MATRICES

In this section we would like to more closely investigate some special simple functions on matrices which are of importance.

First of all, we would like to investigate M^n for $n > N$. This particular function is of interest because the powers of the matrix appear in the developments of the analytic functions over matrices; it is interesting to know, in particular, how the higher powers may be expressed through the fundamental set of powers $(M^0, M^1, \ldots, M^{\bar{N}})$ and invariants; it is also interesting to know how M^n behaves when $n \to \infty$.

According to (11.9)

$$M^n = \sum_{i=1}^{N_0} \sum_{k=0}^{q_i - 1} \binom{n}{k} M_i^{`\,n-k} E_i^{(k)} \quad ; \tag{12.1}$$

consider it for $n >> \bar{N}$. The dependence on n only enters in $\binom{n}{k}$ and in $M_i^{`\,n-k}$. Let $M^{`}_{max}$ be the dominant eigenvalue defined by $| M^{`}_{max} | \geqslant M^{`}_i$. Formula (12.1) makes it obvious that

$$\lim_{n \to \infty} \left(\frac{M}{| M^{`}_{max} | + \epsilon} \right)^n = 0 \quad \epsilon > 0 \, . \tag{12.2}$$

In order to be able to use (12.1) we must know the eigenvalues $M_i^{\backprime}$. It is interesting to observe, however, that directly using (10.2b) we have

$$M^n = \sum_{p=0}^{\bar{N}-1} \left\{ \frac{1}{2\pi i} \oint_{C_\infty} dz\, z^n \frac{\bar{D}_{\bar{N}-1-p}(z)}{\bar{D}_{\bar{N}}(z)} \right\} M^p = \sum_{p=0}^{\bar{N}-1} \bar{C}_p^n \cdot M^p. \tag{12.3}$$

[Because z^n has no singularities, one may freely deform the C of (10.2b) into $C\infty$]. The parallel formula holds when one constructs the resolvent on the basis of $D_N(z)$ instead of $\bar{D}_{\bar{N}}(z)$; it has essentially the same structure as (12.3) but with all the bars above the given symbols omitted. The invariant coefficients $\bar{C}_p^n$ (or C_p^n when one works with $D_N(z)$) may conveniently be expressed in terms of fundamental invariants. Indeed, substitute 'z goes to $1/z$' in the integrals for C_p^n. One obtains

$$C_p^n = \frac{1}{2\pi i} \oint_{C_0} \frac{dz}{z^{(n-p)+1}} \cdot \frac{\psi_{N-1-p}(z)}{\psi_N(z)} \tag{12.4}$$

where the polynomials $\psi_s(z)$ are as defined in (6.3), This formula at once tells s that for $p > n$, C_p^n vanish; for $n \geq p$ we get

$$C_p^n = \frac{1}{(n-p)!} \left(\frac{d}{dz}\right)^{n-p} \frac{\psi_{N-1-p}(z)}{\psi_N(z)} \Bigg|_{z=0} \tag{12.5}$$

This is, however, not the best form obtainable for these coefficients. Indeed, observe that according to (6.3) the following is true

$$\psi_{N-1-p}(z) \equiv \psi_N(z) + (-1)^{N-1-p} z^{N-p} \sum_{q=0}^{p} M_{[N-p+q]} (-z)^q .$$

(12.6)

Therefore, denoting

$$\phi_p(z) = \sum_{q=0}^{p} M_{[N-p+q]} (-z)^q$$

(12.7)

and substituting (12.6) into (12.4) we obtain

$$C_p^n = \delta_p^n + \frac{(-1)^{N-1-p}}{2\pi i} \oint_{C_0} \frac{dz}{z^{(n-N)+1}} \frac{\phi_p(z)}{\psi_N(z)}$$

(12.8)

which implies that

$$0 \leq n \leq N-1 \rightarrow C_p^n = \delta_p^n \qquad (12.9a)$$

$$n \geq N \rightarrow C_p^n = \frac{(-1)^{N-1-p}}{(n-N)!} \left. \frac{d}{dz}^{n-N} \frac{\phi_p(z)}{\psi_N(z)} \right|_{z=0}$$

(12.9b)

In particular, for $n = N$ we obtain

$$C_p^N = (-1)^{N-1-p} \underset{[N-p]}{M}$$

(12.10)

with these values (12.3) [without bars !] reduces to $D_N\left(M\right) = 0$, as should be. (12.9b) gives us the explicit expression for C_p^n in terms of $\underset{[1]}{M}, \ldots, \underset{[N]}{M}$.

Therefore, in order to reduce M^n to lower powers according to $M^n = \sum_{p=0}^{N-1} C_p^n M^p$ we do not need to know the $\grave{M}_i$'s; it is enough to know the $\underset{[p]}{M}$'s. Of course, exactly the same argument, formulae (12.4) through (12.9), may be repeated on the basis of $\bar{D}_{\bar{N}}(z)$; the invariant coefficients of the minimal polynomial cannot, however, be expressed through the fundamental sequences of invariants (or $\underset{[p]}{M}$'s or $\overset{p}{M}$'s). Therefore, such a construction has no practical importance.

Now, a few remarks about $e^M = M\,[e^z]$. This entire function has no singularities (for finite z) and e^M is perfectly well-defined for any M. According to (11.9) we may represent e^M as

$$e^M = \sum_{i=1}^{N_0} e^{\grave{M}_i} \sum_{k=0}^{q_i-1} \frac{1}{k!} E_i^{(k)} \qquad (12.11)$$

When the eigenvalues, multiplicities and their defects are known, the matrices $E_i^{(k)}$ also are explicitly given and formula (12.11) is very convenient when one has to solve a system of the differential equations of the type $\frac{d}{dt} v = M \cdot v$ with constant M.

Up to now we discussed only the one-valued functions on M. The question arises whether one can extend our theory on multi-valued functions and define sensibly quantities as M^s (where s a real number) $\ln M$, etc.

The key to this problem obviously consists in a convenient definition of $\ln M$; having such a matrix one may even define the complex powers of M as: $M^{\alpha} = \exp[\alpha \ln M]$. It is, of course, necessary to demand that $\ln\{F(M) \cdot G(M)\} = \ln F(M) + \ln G(M)$. It follows that in order to approach the problem of $\ln M$ first one has to settle the question of what one should understand by the symbol $\mu = \ln \mathbf{1}$. It is obvious that one has to demand of this matrix: $e^{\mu} = \mathbf{1}$. On the other hand, according to the general theory

$$e^{\mu} = \sum_{i=1}^{N_0} e^{\mu'} \sum_{k=0}^{q_i - 1} \frac{1}{k!} E_i^{(k)} \; (= \mathbf{1})$$

where the $E_i^{(k)}$ are constructed from μ. From this equation we easily conclude (because of the algebra of $E_i^{(k)}$) that: 1. all q_i have to be equal to one, 2. that $e^{\mu'}$ = one. This information gives us a clear insight into possible 'discrete' structures of matrices which may serve as $\ln \mathbf{1}$. A matrix of this type, $\mu \ln \mathbf{1}$, may have $N_0 \geq 1$ *different* eigenvalues $M'_s = 2\pi i\, k_s$, $s = 1,2,\ldots,N_0$, k_s are integers and are different. These may have some multiplicities, n_s, such that $\sum_{s=1}^{N_0} n_s = N$. Because, however, $q_i = 1$, $i = 1,2,\ldots,N_0$ the 'subtle' structure of each of the Jordan's cells is trivial. Consequently, the canonical form of $\ln \mathbf{1}$ is

$$\mu = \ln \mathbf{1} = \left\| \sum_{s=1}^{N_0} 2\pi i \; k_s \sum_{p=1}^{n_s} e^a_{(s,p)} \; \overset{(s,p)}{e_b} \right\| \tag{12.12}$$

where $e_{(s,p)}$, for $s = 1, \ldots, N_0$ and $p = 1, \ldots, n_s$, form a basis; $\overset{(s,p)}{e}$ is its inverse basis. The arbitrariness here is in: 1. integers n_s and k_s, 2. in the choice of the basis. This shows how the usual 'branches' of the logarithm, in a sense, 'multiply themselves' on the level of matrices; there is a lot of arbitrariness in (12,12). When one deals with some definite matrix M which is not proportional to $\mathbf{1}$, one can restrict the matrix which should serve as $\ln \mathbf{1}$ by the additional condition that it should commute with M; when M is maximally undegenerate (N different eigenvalues) we have to pick up in (12.12) $N_0 = N, n_s = 1$, the basis has to be identical with eigenvectors of M; only the k_s here remain arbitrary. But when any degeneration is present even the condition $\mathsf{M} \cdot \ln \mathbf{1} - \ln \mathbf{1} \cdot \mathsf{M} = 0$ leaves the arbitrariness in the basis and in n_s which enter in (12.12).

From these remarks it is clear that in principle one can extend the theory of $F(\mathsf{M})$ on the multi-valued $F(z)$, carefully taking into account all the arbitrarinesses which may occur. To go into the details of such developments, however, lies outside the scope of this paper.

Now, it is also of interest to say a few words about the structure of

$\left(\frac{1}{z_0 - \mathsf{M}}\right)^n = \mathsf{M}\left[\left(\frac{1}{z_0 - z}\right)^n\right]$, the matrix which enters in the principal parts of of the quasi-Laurent development (10.27).

According to the general (11.9) we have

$$\left[\frac{1}{z_0 - M}\right]^n = \sum_{i=1}^{N_0} \sum_{k=0}^{q_i - k} \binom{n-1+k}{n-1} (z_0 - M_i^{`})^{-n-k} \; E_i^{(k)} \tag{12.13}$$

which explains the asymptotic behavior of $\left[z_0 - M\right]^{-n}$ for large n's. In particular, for $n = 1$.

$$\frac{1}{z_0 - M} = \sum_{i=1}^{N_0} \sum_{k=0}^{q_i - 1} \frac{E_i^{(k)}}{\left(z_0 - M_i^{`}\right)^{k+1}} \tag{12.14}$$

This last formula gives us the interpretation of the matrices $E_i^{(k)}$ as just coefficients in the decomposition of $\left[z_0 - M\right]^{-1}$ in simple ratios with respect to $z_0 - M_i^{`}$.

It may be also observed that because of

$$\left[\frac{1}{z_0 - M}\right]^n = \sum_{p=0}^{\bar{N}-1} \left\{ \frac{1}{2\pi i} \oint_{C_\infty} dz \, \frac{\bar{D}_{\bar{N}-1-p}(z)}{(z_0 - z)^n \, \bar{D}_{\bar{N}}(z)} - \frac{1}{2\pi i} \oint_{C_{z_0}} \frac{dz}{(z_0 - z)^n} \, \frac{\bar{D}_{\bar{N}-1-p}(z)}{\bar{D}_{\bar{N}}(z)} \right\} \cdot M^p \tag{12.15}$$

(clearly the contour $C_\infty - C_{z_0}$ is equivalent to C around the singularities of the resolvent), and the fact that the integral over C_∞ vanishes (for $n \geq 1$; this may be seen by the change of variable $z \to 1/z$), we may write

$$\left[\frac{1}{z_0 - \mathsf{M}}\right]^n = \frac{(-)^{n-1}}{(n-1)!} \left(\frac{d}{dz}\right)^{n-1} \sum_{p=0}^{\bar{N}-1} \frac{\bar{D}_{\bar{N}-1-p}(z)}{\bar{D}_{\bar{N}}(z)} \mathsf{M}^p \Bigg|_{z_0} =$$

$$= \frac{(-)^{n-1}}{(n-1)!} \left(\frac{d}{dz_0}\right)^{n-1} \left\{\frac{1}{z_0 - \mathsf{M}}\right\} . \qquad (12.16)$$

When one does this construction on the basis of $D_N(z)$ instead of $\bar{D}_{\bar{N}}(z)$, one gets $\left[z_0 - \mathsf{M}\right]^{-n}$ as the polynomial of $(N-1)$th degree in M with coefficients expressed through $\underset{[r]}{\mathsf{M}}, \ldots, \underset{[N]}{\mathsf{M}}$ and the number z_0; in order to write it down we do not need to know explicitly the eigenvalues and their multiplicities.

Now, for the last problem in this section we shall return to formula (7.32); we were claiming that if the matrix M is of the rank r but $\underset{[r]}{\mathsf{M}}$ is different from zero, then the equations

$$y = \mathsf{M} x \qquad (12.17)$$

have as the consistency condition $\mathfrak{m}_{[r]} \cdot y = 0$, and their solution is given as

$$x = \frac{1}{M_{[r]}} \mathfrak{M}_{[r-1]} \cdot y + \mathfrak{M}_{[r]} \cdot C \tag{12.18}$$

where C is an arbitrary vector; this arbitrariness exhausts all of the possible arbitrariness *in the solution of* (12.17). *Now is the time to clarify the last assertion,* with the help of tools from the previous sections. First of all, rewrite (12.18) as was done in (7.19):

$$x = - \frac{D_{r-1}(M)}{D_r(0)} \cdot y + D_r(M) \cdot C \quad . \tag{12.19}$$

Now, in our case we have $MD_r(M) = 0$ and the characteristic polynomial has the form $D_N(\lambda) = \lambda^{N-r} D_r(\lambda) = (\lambda - 0)^{N-r} \prod_{i=2}^{N_0} \left(\lambda - M_i^{`}\right)^{n_i}$, i.e., we have the eigenvalue $M_1^{`} = 0$ of multiplicity exactly $N - r$ and all the other eigenvalues $M_2^{`}, \ldots, M_{N_0}^{`}$ are necessarily different from zero. It follows that

$D_r(M) = \prod_{i=2}^{N_0} \left(M - M_i^{`}\right)^{n_i}$. This makes it obvious that $D_r(M)$ acting on an arbitrary vector C 'kills' exactly that part of it which cannot be expressed as the linear combination of the eigenvectors belonging to the eigenvalue $M_1^{`} = 0$; hence $D_r(M)$ C is exactly a linear combination of these $N - r$ eigenvectors. Now, as far as the term with y in (12.19) is concerned the necessity that

$\mathfrak{M}_{[r]} \cdot y = (-1)^r D_r(M)\, y = 0$ follows from $MD_r(M) = 0$ and $y = Mx$.
On the other hand, we have the obvious: $D_r(M) - D_r(0) = MD_{r-1}(M)$.
Acting with it on x, which obeys (12.17), we get

$$x = -\frac{D_{r-1}(M)}{D_r(0)} \cdot y + \frac{D_r(M)}{D_r(0)}\, x \ ;$$

here the second term is of the form $D_r(M)$, C, which anyway is arbitrary.
Hence, (12.19) is true.

XII. THE METRIC, TRANSPOSED AND "RECTANGULAR" MATRICES.

The material presented up to now did not require the use of the metric tensor. When, *however, the metric tensor is defined*, i.e., when we have a simple prescription of how to raise (lower) our sets of indices, the tensorial matrices in our sense become richer in properties. The aim of this section is to explore the situation which arises when we have the concept of the metric tensor at our disposal.

Let us return to the considerations of section II. Suppose that we are concerned with *sets* of indices a, b, ... with the internal structure

$$a \equiv \left\{ A'_1, \ldots\ldots, A'_p \;\; ; \;\; A''_1, \ldots\ldots, A''_q \; ; \ldots \right\} \tag{13.1}$$

and with N essentially different values $\overset{k}{\phi}$, k = 1, 2, ..., N. Now, assume that for each kind of internal index there exists a well-defined non-singular metric tensor: $G_{A'B'}$, $G_{A''B''}$, ... The contra-variant metric tensors to these exist and are uniquely defined:

$$G_{A'S'} \; G^{B'S'} = \delta^{B'}_{A'} \quad , \tag{13.2}$$

$$G_{A''S''} \; G^{B''S''} = \delta^{B''}_{A''} \; , \ldots$$

with the help of them we may raise (lower) the internal indices according to the scheme

$$T_{A'} = G_{A'S'} T^{S'}, \quad T_{S'} G^{S'A'} = T^{A'}$$

(similar formulae for A'', etc.). (13.3)

We *do not* assume these metric tensors to be symmetric (e.g., in the case of spinors the metric tensor is skew symmetric) hence it is necessary to preserve the orders of the indices and the contractions as given in (13.2) and (13.3).

Now, with the help of the 'elementary' or 'internal' metric tensors we are able to construct the metric tensors with respect to our *sets* of indices (13.1). Namely, we define

$$G_{ab} = G_{A'_1\{B'_1} \cdots G_{|A'_p|B'_p|} G_{A''_1|B''_1} \cdots G_{|A''_q|B''_q} \cdots \}$$

(13.4)

where the symmetrization of the type specific for our set a is applied here over all B indices. It automatically yields the required internal symmetries for A indices. Quite parallelly, by symmetrizing the external product of contra-invariant G^{AB}'s over, say, B indices we derive G^{ab}. Now, it follows from (13.2) that

$$G_{as} G^{bs} = E_a^b$$

(13.5)

and the rule of raising (lowering) the sets a,b,... which is consistent with (13.3) must be

$$T_a = G_{as} \, T^s , \; T^a = T_s \, G^{sa} \; . \qquad (13.6)$$

Sometimes, when it is necessary to indicate that the given set of indices has been lowered (raised) we will use the notation

$$T_a^{\bullet} = G_{as} \, T^s \; , \; T^a_{\bullet} = T_s \, G^{sa} \; .$$

Because G_{ab} and G^{ab} have both sets of indices on the same level, they are not $N \times N$ matrices in the previous sense. The notion of the determinant may, however, be extended to these quantities.

Namely, we define

$$G = \mathrm{Det} \left\| G_{ab} \right\| = \frac{1}{N!} \, \epsilon^{a_1 \ldots a_N} \, \epsilon^{b_1 \ldots b_N} \, G_{a_1 b_1} \cdots G_{a_N b_N} \; . \qquad (13.7a)$$

$$G^{\cdot\cdot} = \mathrm{Det} \left\| G^{ab} \right\| = \frac{1}{N!} \, \epsilon_{a_1 \ldots a_n} \, \epsilon_{b_1 \ldots b_N} \, G^{a_1 b_1} \ldots G^{a_N b_N} \qquad (13.7b)$$

$$G^{\cdot\cdot} \cdot G_{\cdot\cdot} = 1 \quad . \tag{13.8}$$

Moreover:

$$G^{ab} = \frac{1}{G_{\cdot\cdot}} \cdot \frac{1}{(N-1)!} \epsilon^{a\ a_2 \ldots a_N} \epsilon^{b\ b_2 \ldots b_N} G_{a_2 b_2} \ldots G_{a_N b_N} \quad , \tag{13.9}$$

which justifies the term 'the determinant' as the name for $G_{\cdot\cdot}$. One can prove that $G_{\cdot\cdot}$ is just the product of the determinants (in the literal sense) of the 'internal' metric tensors – of course in powers corresponding to the number of times each given type of the index occurs in the set α .

Now, having the notion of the metric tensor at our disposal, we may introduce the notion of the scalar product in V^N:

$$(U, V) = G_{ab}\, U^a V^b = U_s^{\bullet} V^s \quad (\neq V_s^{\bullet}\, U^s \text{, in general}) \ . \tag{13.10}$$

Moreover, to the given N×N matrix M^a_b we may define the *transposed* matrix $M^{T a}_{b}$ (with the help of the metric) by

$$M^T = \left\| M_{b\bullet}^{\bullet a} \right\| = \left\| G_{bs}\ M^s_r\ G^{ra} \right\| \quad . \tag{13.11}$$

With the help of the notion of transposition we are able now to define a few special types of matrices: symmetric, skew, orthogonal and Hermitian matrices (with respect to the given metric). These we define by demanding

M is *symmetric*	$M = M^T$
M is *skew*	$M = -M^T$
M is *orthogonal*	$M = [M^T]^{-1}$
M is *Hermitian*	$M^* = M^T$
M is *unitary*	$M^* = M^{-1}$

(13.12)

By M^* we understand the complex conjugate matrix which one obtains by applying the complex conjugation to all its elements. Note that in order to be able to define the unitary matrix we did not need the notion of transposition.

The notion of transposition is manitestly covariant because whatever the groups and their representations which govern the internal indices are, the concept of the metric is from definition covariant. Hence, the concepts of symmetric, skew and orthogonal matrix are covariant when the groups and their representations still are very general.

With the concepts of the Hermitian and unitary matrices which use the notion of complex conjugation, however, one has to be very careful in our formalism. M^* forms an object with the same transformation properties as M only with the proviso that the complex conjugation of the transformations acting on its internal indices (with symmetries taken into account) yields the same transformation as these which act on M. This clearly is so when the representations of the groups we have to deal with happen to be real. But, in the case of some specific symmetries, this may also happen to be true in the case of complex representations.

It is important to stress that the consequences of the properties (13.12) become strong and important when the metric G_{ab} has definite simple symmetries in sets a, b, i.e., when it is symmetric: $G_{ab} = G_{ba}$, when it is skew: $G_{ab} = - G_{ba}$, when it is Hermitian: $G^{*}_{ab} = G_{ba}$.

One easily sees that

$$G_{ab} = G_{ba} \text{ implies } \left[M^{P} \right]^{T} = \left[M^{T} \right]^{P} \tag{13.13a}$$

$$G_{ab} = - G_{ba} \text{ implies } \left[M^{P} \right]^{T} = (-1)^{P-1} \left[M^{T} \right]^{P} . \tag{13.13b}$$

Because the properties of the proper powers of the matrix are the key to its structure, as we learned in previous sections, it is clear that relations (13.13) are crucial in the analysis of the consequences of (13.12).

Up to now we delat with 'square' N×N matrices. There are, however, some algebraic problems which also require the use of the concept of the 'rectangular' $\widetilde{N}$×N matrices.

Let

$$a \underset{df}{=} \left\{ A'_{1} \ldots A'_{p} , A''_{1} , \ldots A''_{q} , \ldots \right\} , \widetilde{a} = \left\{ A'_{1} \ldots A'_{\widetilde{p}} , A''_{1} \ldots A''_{\widetilde{q}} , \ldots \right\} \tag{13.14}$$

be two *different* sets of indices, with generally *different* numbers of essentially different values, N, $\widetilde{N}$, and *different* internal symmetries.

The set of complex numbers:

$$\left\| M = M_{b}^{\tilde{a}} \right\| \tag{13.15}$$

we shall call the rectangular $\tilde{N} \times N$ matrix.

Its transposed matrix we define as

$$M^{T} = \left\| M_{b\bullet}^{\bullet a} \right\| = \left\| M_{r}^{\tilde{s}} \; G_{\tilde{b}\tilde{s}} \; G^{ra} \right\| \tag{13.16}$$

and we may understand it as a $N \times \tilde{N}$ matrix.

Now, the products

$$M \cdot M^{T} = \left\| M_{s}^{\tilde{a}} \; M_{\tilde{b}\bullet}^{\bullet s} \right\| , \quad M^{T} \cdot M = \left\| M_{\tilde{s}\bullet}^{\bullet a} \; M_{b}^{\tilde{s}} \right\| \quad , \tag{13.17}$$

are already square $\tilde{N} \times \tilde{N}$ and $N \times N$ matrices to which all are the previous analysis applies. They are crucial in the study of the structure of the rectangular matrices.

Now, a few words about the motivations behind these definitions. The general fields, which one meets in theoretical physics and whose algebraic structure one intends to explore, are usually objects with indices for which metric tensors are well-defined and which therefore may be written with all indices on the same level:

$$\psi_{A'_1 \ldots A'_p \, ; \, A''_1 \ldots A''_q \, ; \ldots} \tag{13.18}$$

These indices may have some symmetries or, in particular, none. Now, picking up some of these indices so that they will form a set a, raising them by the help of the metric, calling the remaining set $\tilde{a}$, we may understand our object as

$$\Psi^{\tilde{a}}{}_{b} \,, \tag{13.19}$$

i.e., a $\tilde{N} \times N$ rectangular matrix. If it is possible to thus pick up the set $\tilde{a}$, that which remains forms a set $b \equiv \tilde{b}$, with the *same* internal structure, we may understand $\Psi^{\tilde{a}}{}_{\tilde{b}}$ as a $\tilde{N} \times \tilde{N}$ matrix to which all our previous analysis applies. There are, however, some objects where such grouping of indices is impossible (e.g., Ψ_{ABC}, a spinorial field with an odd number of indices). The best that one can do, if one wants to apply the methods of the linear algebra, is to form from such general object a rectangular $\tilde{N} \times N$ matrix and to investigate its structure. This explains why the study of rectangular matrices is important. Note that a general object like (13.18) may be put into the form (13.19) in many ways – by making all possible choices for the set a from the original sequency of indices (and making $\tilde{a}$ contain one or two, or three etc. of the internal indices). It is intuitively obvious that the algebraic characteristics of all possible rectangular matrices which one can form from (13.18) must completely describe the structure of the original object.

It is my pleasure to express my gratitude to Dr. Arturo Rosenblueth for his warm hospitality at the Centro de Investigación y de Estudios Avanzados del I.P.N.

I owe my gratitude to Mr. Bogdan Mielnik for his very active help in the process of writing this paper. Especially appreciated is his cooperation in all problems concerning nil-potent matrices. Several ideas developed in the sections about nil-potent matrices and canonical forms were proposed by him.

I am also grateful to Mrs. Okhuysen for her kind assistance in the preparation of the manuscript.

REFERENCES

1 A.Z. Petrov, 'Prostranstva Einsteina', Gost. Izd., Moscov (1961)

2 F.A.E. Pirani, Acta Phys. Pol. **15**, 389 (1956) and Phys. Rev., **105**, 1089 (1957)

3 P. Jordan, J. Ehlers, R.K. Sachs, Akad. Wiss. Mainz Abh. Math.- nat. Kl. 1960, Nr. 2 and 1961, Nr. 1

4 R. Sachs, Proc. Roy. Soc. (A) **266**, 309 (1961), **270**, 103 (1962)

5 R. Debever, C.R. Acad. Sci. Paris **249**, 1324, 1744 (1958)

6 R. Penrose, Null Hypersurface Initial Data . . . (Preprint)

7 R. Penrose, Zero Rest Mass Fields Including Gravitation . . . (Preprint)

8 R. Penrose, Annals of Phys., **10**, 171 (1960)

9 I. Robinson and A. Schild, J. Math. Phys., **4**, 484 (1963)