Bogue

NATIONAL BUREAU OF STANDARDS REPORT

2007

FOUR ARTICLES ON NUMERICAL MATRIX METHODS

Translated from the Russian by Curtis D. Benster

Editor: G. E. Forsythe
National Bureau of Standards

NBS

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

U. S. DEPARTMENT OF COMMERCE
Charles Sawyer, *Secretary*

NATIONAL BUREAU OF STANDARDS
A. V. Astin, [redacted] *Director*

THE NATIONAL BUREAU OF STANDARDS

The scope of activities of the National Bureau of Standards is suggested in the following listing of the divisions and sections engaged in technical work. In general, each section is engaged in specialized research, development, and engineering in the field indicated by its title. A brief description of the activities, and of the resultant reports and publications, appears on the inside of the back cover of this report.

1. **Electricity.** Resistance Measurements. Inductance and Capacitance. Electrical Instruments. Magnetic Measurements. Electrochemistry.

2. **Optics and Metrology.** Photometry and Colorimetry. Optical Instruments. Photographic Technology. Length. Gage.

3. **Heat and Power.** Temperature Measurements. Thermodynamics. Cryogenics. Engines and Lubrication. Engine Fuels.

4. **Atomic and Radiation Physics.** Spectroscopy. Radiometry. Mass Spectrometry. Physical Electronics. Electron Physics. Atomic Physics. Neutron Measurements. Nuclear Physics. Radioactivity. X-Rays. Betatron. Nucleonic Instrumentation. Radiological Equipment. Atomic Energy Commission Instruments Branch.

5. **Chemistry.** Organic Coatings. Surface Chemistry. Organic Chemistry. Analytical Chemistry. Inorganic Chemistry. Electrodeposition. Gas Chemistry. Physical Chemistry. Thermochemistry. Spectrochemistry. Pure Substances.

6. **Mechanics.** Sound. Mechanical Instruments. Aerodynamics. Engineering Mechanics. Hydraulics. Mass. Capacity, Density, and Fluid Meters.

7. **Organic and Fibrous Materials.** Rubber. Textiles. Paper. Leather. Testing and Specifications. Organic Plastics. Dental Research.

8. **Metallurgy.** Thermal Metallurgy. Chemical Metallurgy. Mechanical Metallurgy. Corrosion.

9. **Mineral Products.** Porcelain and Pottery. Glass. Refractories. Enameled Metals. Building Stone. Concreting Materials. Constitution and Microstructure. Chemistry of Mineral Products.

10. **Building Technology.** Structural Engineering. Fire Protection. Heating and Air Conditioning. Exterior and Interior Coverings. Codes and Specifications.

11. **Applied Mathematics.** Numerical Analysis. Computation. Statistical Engineering. Machine Development.

12. **Electronics.** Engineering Electronics. Electron Tubes. Electronic Computers. Electronic Instrumentation.

13. **Ordnance Development.** Mechanical Research and Development. Electromechanical Fuzes. Technical Services. Missile Fuzing Research. Missile Fuzing Development. Projectile Fuzes. Ordnance Components. Ordnance Tests. Ordnance Research.

14. **Radio Propagation.** Upper Atmosphere Research. Ionospheric Research. Regular Propagation Services. Frequency Utilization Research. Tropospheric Propagation Research. High Frequency Standards. Microwave Standards.

15. **Missile Development.** Missile Engineering. Missile Dynamics. Missile Intelligence. Missile Instrumentation. Technical Services. Combustion.

NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT — 1101-10-5100

August 14, 1952

NBS REPORT — 2007

FOUR ARTICLES ON NUMERICAL MATRIX METHODS

Translated from the Russian by Curtis D. Benster*

Editor: G. E. Forsythe
National Bureau of Standards

PREPRINT

*These translations were sponsored in part by the Office of Naval Research

NBS

CONTENTS

I.

A NUMERICAL METHOD FOR DETERMINING THE CHARACTERISTIC VALUES AND CHARACTERISTIC PLANES OF A LINEAR OPERATOR

by

A. M. Lopshits [1]

The numerical method of determining the coefficients of the characteristic equation of a linear operator (the characteristic equation of a matrix) that was suggested in 1931 by Academician A. N. Krylov [1] required a considerably smaller quantity of computations than the methods that had been developed earlier. Nevertheless neither this method nor that published in 1937 by A. Danilevsky [2] (which reduces the numerical work to approximately two-thirds of that required by the Krylov method) effected simplifications in the solution of the problem of determining the characteristic vectors of a linear operator. A geometrical method that I have suggested [3], [4], which leads to the construction of

- - - - - - - - - - - - - -

[1] Translated from Moscow, universitet, fiziko-mekhanicheskiĭ fakul'tet, nauchno-issledovatel'skiĭ institut matematiki i mekhaniki, seminar po vektornomu i tenzornomu analizu ..., Trudy, vol. 7 (1949), pp. 233-259.

Translator's acknowledgement: In construing two passages by which he was balked, the translator has enjoyed the assistance of Prof. Dimitry M. Krassovsky, Slavic Consultant, UCLA, to whom his thanks are extended.

the characteristic equation in the form in which it is suggested by Academician A. N. Krylov, offers the possibility, with a suitable continuation that is expounded in the present article, of constructing a new algorithm for the solution of the problems indcated in the title. This algorithm provides a new geometrical scheme for reduction of the matrix to Jordan normal form.

§1. If in the sequence of vectors

$$a, Aa, A^2a, \dots$$ [1]

only the first m are linearly independent, we then have the equation

$$A^m a + \alpha_1 A^{m-1} a + \alpha_2 A^{m-2} a + \dots + \alpha_{m-1} Aa + \alpha_m a = 0 \quad , \tag{1}$$

where the numerical coefficients $\alpha_1, \dots, \alpha_m$ are uniquely defined in terms of the initial vector a (for a given operator A; a discussion of a computational scheme making possible the determination of the coefficients α_i in terms of the coordinates of the vectors a, Aa, $A^2a, \dots, A^{m-1}a, A^m a$ is given in §9).

The plane[2] $\mathfrak{A}$, defined by the vectors $a, Aa, \dots, A^{m-1}a$

---

[1]The Roman-type minuscule a designates a vector of an n-dimensional vector space; the Roman majuscule A designates a linear vector function of a vector argument (a linear operator) relating to the vector a the vector Aa. The product of the operators A and B, i.e., the operator C, is defined by the equation $Cx = A(Bx)$, and we shall write $C \equiv AB$. Let us agree also on the conventions $AA = A^2$; $AA^2 = A^3$, etc.

[2]That is, the manifold of all vectors that are linear combinations of the vectors $A^i a$ $(i = 1, 2, \dots, m-1)$.

("belonging" to the vector a) is obviously an invariant plane of the operator A (i.e., contains the vector Ap if the vector p is taken from this same plane). If we introduce into the discussion the "characteristic polynomial $\phi(\lambda)$ belonging to the vector a":

(1') $$\phi(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \cdots + \alpha_{m-1}\lambda + \alpha_m \quad ,$$

it is easily seen that, for any vector x lying in the plane $\mathfrak{A}$, the equation

(2) $$\phi(A)x = 0 \quad ,$$

holds, where

$$\phi(A) = A^m + \alpha_1 A^{m-1} + \cdots + \alpha_{m-1}A + \alpha_m \quad .$$

Indeed, multiplying (1) by A^i, we obtain

$$A^i\phi(A)x = 0$$

or, on the strength of the commutativity of polynomials in the operator A, we obtain

$$\phi(A)A^i a = 0 \ , \qquad i = 0, 1, 2, \cdots, m-1 \quad ,$$

i.e., equation (2) holds for the m linearly independent vectors $A^i a$ defining the plane $\mathfrak{A}$.

§2. Let us employ the designation

(3) $$\frac{\omega(\lambda) - \omega(\lambda_1)}{\lambda - \lambda_1} \equiv D\omega(\lambda) \quad ,$$

$\omega(\lambda)$ being an arbitrary polynomial in λ, and λ_1 a given number; let us also use the designations

(3') $$D^i\omega(\lambda) = DD^{i-1}\omega(\lambda) \quad ; \qquad D^0\omega(\lambda) = \omega(\lambda) \quad .$$

If λ_1 is a root of multiplicity k of the characteristic polynomial $\phi(\lambda)$, we obviously have

$$D^i\phi(\lambda) = \frac{\phi(\lambda)}{(\lambda - \lambda_1)^i} \quad (i = 1, 2, \cdots, k) \quad . \tag{3''}$$

Let us introduce into the discussion the "structural vectors" $a_1, a_2, \cdots, a_k$, defined by the equations

$$a_i = D^i\phi(A)a \quad (i = 1, 2, \cdots, k) \quad , \tag{4}$$

and which "belong to the root λ_1".

It is easily shown that the structural vectors a_i are linearly independent and satisfy the system of equations

$$(A - \lambda_1)a_i = a_{i-1} ; \quad i = 1, 2, \cdots, k; \quad a_o = 0 \quad . \tag{5}$$

Indeed the vector a_i is, on the strength of formula (4), a linear combination of the linearly independent vectors a, Aa, $\cdots$, $A^{m-i}a$, in which the coefficient of the vector A^{m-i} is equal to unity, whence it then follows that the vector a_i is linearly independent of the vectors $a_1, a_2, \cdots, a_{i-1}$. Equations (5) are, moreover, a consequence of the identity

$$(A - \lambda_1)D^i\phi(A) = D^{i-1}\phi(A) \quad .$$

If λ_1 is a root of multiplicity one, we obtain a unique structural vector belonging to the root λ_1 and satisfying (on the strength of (5)) the equation

$$Aa_1 = \lambda_1 a_1 \quad ,$$

i.e., we obtain the so-called characteristic vector <u>a</u> of the operator A and the characteristic number (characteristic value), λ_1, that

corresponds to it. The computation of the coordinates of this characteristic vector reduces, therefore, in accordance with formula (3"), to the computation of the coordinates of the vector

$$a_1 = \frac{\phi(A)}{A - \lambda_1} a \equiv A^{m-1}a + \beta_1 A^{m-2}a + \cdots + \beta_{m-2}Aa + \beta_{m-1}a \quad ,$$

i.e., to the computation of the coordinates of a linear combination of the given vectors. (A calculation of the number of computational operations required for this, and a comparison with the number of operations required for the solution of this problem by existing methods, will be given below.)

If the multiplicity of the root λ_1 be greater than 1 $(k > 1)$, formulas (5) show that the structural vectors $a_1, a_2, \cdots, a_i$ define an invariant plane belonging to the vector a_i to which the characteristic polynomial $(A - \lambda_1)^i$ belongs. We shall call such a plane a characteristic plane of the operator A (an axial manifold [5], [6]). (Thus the characteristic vector is a "characteristic plane of one dimension.") Formulas (4) give us the possibility of computing the coordinates of the structural vectors $a_1, a_2, \cdots, a_k$ that define the characteristic plane $\mathfrak{a}_1$ belonging to the characteristic number λ_1.

Let us note that for any vector x lying in this plane, the equality

$$(A - \lambda_1)^k x = 0$$

holds.

§3. If the degree of the constructed characteristic polynomial $\phi(\lambda)$ belonging to the vector a is equal to the number of

dimensions of the vector space ($m = n$), then, having constructed in the manner indicated the system of structural vectors for each root λ_i of multiplicity k_i (which vectors define a k_i-dimensional characteristic plane α_i), we arrive at a complete decomposition of the vector space into characteristic planes, since two characteristic planes belonging to different roots will not have a common part.[1] In this case (to which, strictly speaking, A. N. Krylov [1] limited himself), for the determination of the characteristic numbers and characteristic planes the following operations must be effected:

1. The computation of the coordinates of n vectors: Aa, A^2a, $\cdots$, $A^{n-1}a$, A^na.

2. The resolution of the vector A^na in terms of the linearly independent vectors a, Aa, $\cdots$, $A^{n-1}a$ (the computation of the coefficients of the characteristic equation[2] consists of just

- - - - - - - - - - - - - -

[1]See [4].

[2]The determination of these coefficients can be effected, on the strength of (1) (if we consider that $m = n$) by means of the formulas

$$\alpha_i = (a, Aa, \cdots, A^{n-i-1}a, A^{n-i+1}a, \cdots, A^na)\cdot\delta_i^{-1} \quad (i=1,2,\cdots,n),$$

where

$$\delta_i = (-1)^i(a, Aa, \cdots, A^{n-1}a, A^na)$$

(the symbol $(p_1, \cdots, p_n)$ here denotes the determinant whose s-th row is the aggregate of the n successive coordinates of the vector p_s). Substituting these expressions for α_i in formula (1'), we

this), i.e., the solution of a system of n linear equations in n unknowns.[1]

3. The computation of the roots of the characteristic equation.

4. The computation for each root λ_i of multiplicity k_i of the coefficients of the polynomials $\frac{\phi(\lambda)}{(\lambda - \lambda_i)^j}$, $j = 1, 2, \cdots, k_i$, i.e., the computation of the coordinates of the structural vectors of the characteristic plane (belonging to the root λ_i) with respect to the coordinatal system: $a, Aa, \cdots, A^{n-1}a$.

5. The computation of the coordinates of the structural vectors in the initial coordinate system (in accordance with formula (4)).

If all roots of the polynomial are of multiplicity one, for computing the n characteristic numbers and the coordinates of the n corresponding characteristic vectors there will be required (if the operations required for the determination of the roots of the characteristic equation be disregarded) about $4n^3$ additions and multiplications.

To solve this same problem using Danilevsky's method [2] for

- - - - - - - - - - - - -

obtain the characteristic polynomial of the affinor A in the form in which it was proposed in 1931 by Acad. A. N. Krylov. The method given above for obtaining this polynomial was given in 1933 by me [3]. See also [4], p. 176.

[1]See below, §9.

the computation of the coefficients of the characteristic equation, if for each root we compute by the familiar method [6] the coordinates of the corresponding characteristic vector (i.e., solve the corresponding homogeneous system of n linear equations in n unknowns), about $2(n-1)^3 n$ additions and subtractions will be required, i.e., approximately $(n-1)^3/2n^2$ as many operations as does the method referred to above.

§4. If the characteristic equation has the complex root $\lambda_1 = \alpha + i\beta$ of multiplicity k_1, the effectuation of the operations indicated above is complicated computationally. In order to reduce the number of operations a preliminary construction of a system of polynomials with real coefficients

$$\psi_s(\lambda) = \frac{\phi(\lambda)}{(\lambda^2 - 2\lambda\alpha + \alpha^2 + \beta^2)^s} \qquad (s = 1, 2, \cdots, k) \quad ,$$

is called for, as is also the construction of a system of linearly independent real vectors

$$b_s = \psi_s(A)a \qquad (s = 1, 2, \cdots, k) \quad ,$$

which will obviously satisfy the structural equations

$$(A^2 - 2\alpha A + \alpha^2 + \beta^2)b_s = b_{s-1} \qquad (s = 1, \cdots, k_1;\ b_s = 0) \quad .$$

One has no difficulty in showing that the characteristic plane belonging to the complex characteristic number $\alpha + i\beta$ is determined by the (complex) structural vectors

$$a_i = (A - (\alpha - i\beta))^i b_i \qquad (i = 1, 2, \cdots, k_1) \quad .$$

§5. If $m < n$, the characteristic planes constructed in §1 do not exhaust the vector space; for the determination of the new characteristic planes the process indicated above may be repeated, choosing for the initial vector an arbitrary vector $\underline{a}'$ not lying in the plane $\mathfrak{A}$. If the characteristic polynomial $\phi'(\lambda)$ belonging to this vector is of degree n, the decompositon of the space into characteristic planes may be conducted as has been indicated above. The computational labor spent in determining the coefficients of the polynomial $\phi(\lambda)$ proves not to be utterly superfluous here: the polynomial $\phi(\lambda)$ is a divisor of the polynomial $\phi'(\lambda)$ and this fact of course facilitates the computation of the roots of the polynomial $\phi'(\lambda)$.

If, however, the degree n' of the polynomial $\phi'(\lambda)$ is less than n, we will obtain, for each root λ_1' of it that is different from the roots of the polynomial $\phi(\lambda)$, a new characteristic plane that does not have a part in common with the plane $\mathfrak{A}$. If on the other hand a root equal to a root λ_1 of the polynomial $\phi(\lambda)$ is to be found among the roots of the polynomial $\phi'(\lambda)$, let the characteristic plane $\mathfrak{A}'$, defined by the structural vectors

$$a_i' = D^i\phi'(A)a' \qquad (i = 1, 2, \cdots, k) \quad .$$

belong to it.

The relative situation of these two characteristic planes belonging to the same root λ_1 is determined in accordance with the following

Theorem: <u>Let the characteristic planes</u>

$$\mathfrak{a}(a_1, a_2, \cdots, a_k) \text{ and } \mathfrak{a}'(a_1', a_2', \cdots, a_\chi') \quad k \geq \chi$$

belong to the same root λ_1, i.e., let

(I) $(A - \lambda_1)a_i = a_{i-1}$; $a_0 = 0$ $(i = 1, 2, \cdots, k)$,

(II) $(A - \lambda_1)a_j' = a_{j-1}'$; $a_0' = 0$ $(j = 1, 2, \cdots, \chi)$

and, furthermore, let the following equalities hold:

$$\text{(III)} \quad \left.\begin{array}{l} a_1' = \alpha_1 a_1 \quad , \\ a_2' = \alpha_1 a_2 + \alpha_2 a_1 \quad , \\ a_3' = \alpha_1 a_3 + \alpha_2 a_2 + \alpha_3 a_1 \quad , \\ \cdots\cdots\cdots\cdots\cdots \\ a_m' = \alpha_1 a_m + \alpha_2 a_{m-1} + \cdots + \alpha_m a_1 \quad . \end{array}\right\}$$

If, along with this, the vector

$$c_1 \equiv a_{m+1}' - (\alpha_1 a_{m+1} + \alpha_2 a_m + \cdots + \alpha_m a_2)$$

is not collinear with the vector a_1, the vectors

$$c_1, c_2, \cdots, c_{\chi-m} \quad ,$$

defined by the formulas

$$c_i = a_{m+i}' - (\alpha_1 a_{m+i} + \alpha_2 a_{m+i-1} + \cdots + \alpha_m a_{i+1})$$

$$\text{(IV)} \qquad (i = 1, 2, \cdots, \chi - m) \quad ,$$

determine a characteristic plane $\mathcal{L}$ belonging to the root λ_1 and having no part in common with the characteristic plane $\mathfrak{a}$.

Proof. Applying the operator $(A - \lambda)$ to both parts of equation (IV), we obtain, using (I) and (II):

$$(A - \lambda_1)c_i = a'_{m+i-1} - (\alpha_1 a_{m+i-1} + \alpha_2 a_{m+i-2} + \cdots + \alpha_m a_i)$$

and, accordingly (on the strength of (IV) and (III)),

(V) $$(A - \lambda_1)c_i = c_{i-1} ; \quad c_0 = 0 ; \quad i = 1, 2, \cdots, \chi - m \quad .$$

It thus remains to be shown that the vectors $c_1, c_2, \cdots, c_{\chi-m}$ are linearly independent and form a region[1] $\mathcal{L}$ having no common part with region $\mathfrak{A}$.

Let us first satisfy ourselves that the vector c_1 does not lie in region $\mathfrak{A}$. Indeed, in the contrary case the equality

(VI) $$c_1 = \sigma_1 a_1 + \sigma_2 a_2 + \cdots + \sigma_k a_k \quad .$$

would hold.

Applying the operator $(A - \lambda)$ to both sides, we obtain, on the strength of (V) and (I):

$$0 = \sigma_2 a_1 + \sigma_3 a_2 + \cdots + \sigma_k a_{k-1}$$

and consequently

$$\sigma_2 = \sigma_3 = \cdots \sigma_k = 0 \quad .$$

Substituting these values for σ_i in (VI), we would have

$$c_1 = \sigma_1 a_1 \quad ,$$

i.e., we are led to a contradiction with the condition of the Theorem.

We shall now show that the vectors c_1 and c_2 are linearly

- - - - - - - - - - - - -

[1]Editor's note: Here and below the author uses 'region' (oblast') for 'sub-space.'

independent. Indeed, if the equality

$$c_2 = \sigma c_1 \quad ,$$

held, we would have the equation

$$(A - \lambda_1)c_2 = \sigma(A - \lambda_1)c_1 \quad ,$$

or, in accordance with (5),

$$c_1 = 0 \quad ,$$

which contradicts the condition.[1] We shall now show that the plane formed by the vectors c_1 and c_2 does not have a part in common with the plane $\mathfrak{A}$. Indeed, if the equality

$$\alpha_1 c_1 + \alpha_2 c_2 = \sigma_1 a_1 + \sigma_2 a_2 + \cdots + \sigma_k a_k \; ; \qquad a_2 \neq 0 \; ,$$

held, then on applying the operator $(A - \lambda_1)$ to both parts of it we would obtain (on the strength of (V)):

$$\alpha_2 c_1 = \sigma_2 a_1 + \cdots + \sigma_k a_{k-1} \quad ,$$

i.e., we would be led to a contradiction with the already proven thesis that the vector c_1 does not lie in the region $\mathfrak{A}$.

Continuing analogously, we become convinced that the vector c_3 does not lie in the region formed by the vectors c_1 and c_2 (for from equality

$$c_3 = \alpha_1 c_1 + \alpha_2 c_2$$

it would follow that

$$(A - \lambda_1)c_3 = \alpha_1(A - \lambda_1)c_1 + \alpha_2(A - \lambda_1)c_2 \quad ,$$

- - - - - - - - - - - - - - -

[1]By this it is also proved that $c_2 \neq 0$.

i.e., that $c_2 = \alpha_2 c_1$), and we show that the region formed by the vectors c_1, c_2, c_3 does not have a part in common with the region $\mathfrak{A}$, since, applying to both parts of the equation

$$\alpha_1 c_1 + \alpha_2 c_2 + \alpha_3 c_3 = \sigma_1 a_1 + \sigma_2 a_2 + \cdots + \sigma_k a_k \; ; \quad \alpha_3 \neq 0$$

the operator $(A - \lambda)$, we would arrive at the equation

$$\alpha_2 c_1 + \alpha_3 c_3 = \sigma_2 a_1 + \cdots + \sigma_k a_{k-1} , \quad \alpha_3 \neq 0 ,$$

which contradicts the already proven thesis that the plane c_1, c_2 has no common part with the plane $\mathfrak{A}$. The proof can now be completed in the obvious fashion.

Let us remark in conclusion that if $m = 0$ (i.e., if the vectors a_1 and a_1' are not collinear), the regions $\mathfrak{A}$ and $\mathfrak{A}'$ have no part in common and $\mathfrak{A}' = \mathcal{L}$. If, however, $m \neq 0$, region $\mathfrak{A}'$ then lies in region $\mathfrak{A}$ and accordingly $\mathcal{L} = 0$.

Thus for each root λ_i of the polynomial $\phi'(\lambda)$ that is also a root of the polynomial $\phi(\lambda)$ we shall either find a new characteristic plane $\mathcal{L}_i$ belonging to this root and not having a part in common with the characteristic plane $\mathfrak{A}_i$, found earlier (and belonging to this same characteristic number and lying in the plane of the vectors a, Aa, $\cdots$, $A^{m-1}a$), or we shall find a characteristic plane $\mathcal{L}_i$ containing as a part (proper or improper) of itself the plane $\mathfrak{A}_i$ found earlier.

If the aggregate of the new characteristic planes that can be constructed as indicated by means of the polynomial $\phi'(\lambda)$ and of the vectors $A^i a'$ $(i = 0, 1, \cdots, n_1')$ fills out the aggregate of characteristic planes constructed earlier to an n-dimensional

vector space, the solution of the problem has then been completed. If, however, the n-dimensional space is not exhausted with all the characteristic planes that have been constructed, the aggregate of all the characteristic planes constructed will nonetheless be of greater dimension than the dimension of the plane $\mathfrak{A}$, since the vector a' was chosen outside the plane $\mathfrak{A}$. (This dimension may nevertheless be less than the sum of the dimensions of the planes $\mathfrak{A}$ and $\mathfrak{A}'$; this circumstance may be awkwardly reflected in the number of computations necessary to arrive at a complete solution of the problem.) Continuing, therefore, the construction by means of a new initial vector a" that does not lie in the region $\mathfrak{A} + \mathfrak{A}'$, we shall arrive either at an exhaustion of the space or at a further increase of the dimension of the region made up of the constructed characteristic planes. Thus a finite number of steps will lead us to the complete exhaustion of the space by the characteristic planes.

The algorithm proposed above also solves the problem of the reduction of the matrix corresponding to the operator A to Jordan canonical form: one need only write the matrix corresponding to the operator A in the coordinatal system determined by the aggregate of the structural vectors of all the characteristic planes, suitably renumbering them.

§6. It was stated above (§5) that in case the degree of the characteristic polynomial $\phi(\lambda)$ is less than n, it may happen that the subsequent stages of the computations utilize quite incompletely

the computational work carried out on the preceding stages. We will now give a new scheme that fully utilizes at each new step of the computations the results obtained in the preceding stages.

Assuming that $m < n$ and selecting an arbitrary vector a' that does not lie in the plane $\mathfrak{A}$, let us construct a sequence of vectors $A^i a'$ such that the vectors

$$\text{(6)} \qquad a,\ Aa,\ A^2a,\ \cdots,\ A^{m-1}a,\ a',\ Aa',\ \cdots,\ A^{m'-1}a'$$

are linearly independent and such that the following equality holds:

$$(A^{m'}a' + \alpha_1' A^{m'-1}a' + \cdots + \alpha'_{m'-1}Aa' + \alpha'_{m'}a')$$

$$\text{(6')} \qquad + (\beta_1 A^{m-1}a + \beta_2 A^{m-2}a + \cdots + \beta_m a) = 0 \quad .$$

The plane defined by the vectors (6) we shall denote by $\mathfrak{A} + \mathfrak{A}'$. Let us introduce into the discussion the polynomials

$$\phi'(\lambda) = \lambda^{m'} + \alpha_1'\lambda^{m-1} + \cdots + \alpha_m' \quad ;$$

$$\psi(\lambda) = \beta_1\lambda^{m-1} + \beta_2\lambda^{m-2} + \cdots + \beta_{m-1}\lambda + \beta_m$$

and rewrite equation (6') in the form

$$\text{(6")} \qquad \phi(A)a' + \psi(A)a = 0 \quad .$$

Let λ_1 be a root of the polynomial $\phi'(\lambda)$ of multiplicity k' and at the same time a root of the polynomial $\phi(\lambda)$ of multiplicity k. We shall now show a method of constructing the system of vectors

$$a_1',\ a_2',\ \cdots,\ a_k'$$

and the system of numbers

$$\sigma_0,\ \sigma_1,\ \cdots,\ \sigma_{k'-1} \quad ,$$

satisfying the equations

(7) $(A - \lambda_1)a'_i = a'_{i-1} + \sigma_{i-1}a_k$, $i = 1, 2, \cdots, k'$; $a'_o = 0$,

where a_k is the last structural vector of the characteristic plane; it lies in the plane $\mathfrak{A}$ (defined in §2) and belongs to the characteristic number λ. If $k = 0$, i.e., if λ_1 is not a root of the polynomial $\phi(\lambda)$, then $a_k = 0$.

With the aforementioned objective in view, let us rewrite equation (6") in the form

(8) $$(A - \lambda_1)\,(D\phi'(A)a' + D\psi(A)a) + \psi(\lambda_1)a = 0 \quad ,$$

where the operator D is defined by formula (3).

Taking into account that, in accordance with (4),

$$a_k = \frac{\phi(A)}{(A - \lambda_1)^k}a = D^k\phi(A)a \quad ,$$

we obtain

$$a_k = (A - \lambda_1)D^{k+1}\phi(A)a + D^k\phi(\lambda_1)a$$ [1]

and accordingly

(9) $$a = \frac{1}{D^k\phi(\lambda_1)}a_k - (A - \lambda_1)\,\frac{D^{k+1}\phi(A)}{D^k\phi(\lambda_1)}a$$

(since $D^k\phi(\lambda_1) \neq 0$). Substituting the expression thus obtained for a in equation (8), we obtain:

$$(A - \lambda_1)\left(D\phi'(A)a' + D\psi(A)a - \psi(\lambda_1)\,\frac{D^{k+1}\phi(A)}{D^k\phi(\lambda_1)}a\right) + \frac{\psi(\lambda_1)}{D^k\phi(\lambda_1)}a_k = 0 \; .$$

- - - - - - - - - - - - -

[1] Here $D^k\phi(\lambda_1) \equiv (D^k\phi)(\lambda_1) \equiv (D^k\phi(\lambda))_{\lambda=\lambda_1}$.

Let us introduce into the discussion the operator P, defined by the equation

$$P\omega(\lambda) = \omega(\lambda) - \omega(\lambda_1)\,\frac{D^k\phi(\lambda)}{D^k\phi(\lambda_1)} \quad ;$$

introducing also the designations:

$$a_1' = D\phi'(A)a' + DP\psi(A)a \tag{10}$$

and

$$-\sigma_0 = \frac{\psi(\lambda_1)}{D^k\phi(\lambda_1)} \quad ,$$

we arrive at the result:

$$(A - \lambda)a_1' = \sigma_0 a_k \quad ,$$

i.e., at the first of the equations of system (7).

Furthermore, rewriting equation (10) in the form:

$$a_1' = (A - \lambda_1)(D^2\phi'(A)a' + D^2P\psi(A)a) + DP\psi(\lambda_1)a$$

and substituting for a in the second summand of the right member of (10) its expression (9), we obtain the second equation of system (7):

$$(A - \lambda_1)a_2' = a_1' + \sigma_1 a_k \quad ,$$

where

$$a_2' = D^2\phi'(A)a' + (DP)^2\psi(A)a \quad ,$$

$$-\sigma_1 = \frac{DP\psi(\lambda_1)}{D^k\phi(\lambda_1)} \quad .$$

Continuing in analogous fashion, we will also obtain the subsequent equations of system (7), but for the vectors a_i' and the scalars σ_i' we have the following expressions:

$$a_i' = D^i\phi'(A)a' + (DP)^i\psi(A)a \qquad (i = 1, 2, \cdots, k') \quad ,$$

(10')

$$-\sigma_i = \frac{(DP)^i\psi(\lambda_1)}{D^k\phi(\lambda_1)} \qquad (i = 0, 1, \cdots, k'-1) \quad .$$

If the vector $a_k = 0$ (i.e., if λ_1 is not a root of the polynomial $\phi(\lambda)$) all the numbers $\sigma_o, \cdots, \sigma_{k'-1}$ are equal to zero, and the vectors

$$a_1', a_2', \cdots, a_{k'}'$$

are (in accordance with (7)) structural vectors of the characteristic plane $\mathfrak{A}_1'$ having no part in common with the characteristic plane $\mathfrak{A}_1$ constructed earlier.

If, however,

$$a_k \neq 0$$

and

$$\sigma_o = \cdots = \sigma_{\chi-1} = 0 \; ; \qquad \sigma_\chi \neq 0 \;, \quad 0 \leq \chi < k' \quad ,$$

then we construct the system of vectors b_i defined by the recurrence relations:

(11) $b_{i-1} = (A - \lambda_1)b_i \qquad (i = k'-1, k'-2, \cdots) \; ; \quad b_{k'} = a_{k'}' \quad .$

Utilizing (7) and (5), we obtain

$$b_i = a_i' + \sigma_i a_k + \sigma_{i+1}a_{k-1} + \cdots + \sigma_{k'-1}a_{k-(k'-i+1)}$$

and, in particular,

$$b_1 = a_i' + \sigma_1 a_k + \sigma_2 a_{k-1} + \cdots$$

(here and henceforth it is assumed that $a_s = 0$ if $s < 0$). Therefore

$$b_o = (A - \lambda_1)b_1 = \sigma_o a_k + \sigma_1 a_{k-1} + \sigma_2 a_{k-2} + \cdots$$

and accordingly

$$b_0 = (A - \lambda_1)b_1 = \sigma_\chi a_{k-\chi} + \sigma_{\chi+1}a_{k-\chi-1} + \cdots \quad .$$

If $k - \chi \leqq 0$,

$$b_1 = a_1' + \sigma_\chi a_{k-\chi+1} \quad ,$$

$$(A - \lambda_1)b_1 = 0$$

and consequently the constructed (linearly independent) vectors $b_1, b_2, \cdots, b_k$, are structural vectors of the characteristic plane $\mathfrak{A}_1'$ which has no part in common with the plane $\mathfrak{A}_1$ constructed earlier (since the vector b_1 is not collinear with the vector a_1').

If $k - \chi > 0$, however, then $b_0 \neq 0$ and, continuing the construction prescribed by the recurrence formula (11), we construct the vectors

$$b_{-1};\ b_{-2};\ \cdots;\ b_{(k-\chi)+1} = \sigma_\chi a_1 \quad ;$$

the (linearly independent) vectors

$$b_{(k-\chi)+1};\ b_{(k-\chi)+2};\ \cdots ,\quad b_0;\ b_1;\ b_2;\ \cdots ,\quad b_k$$

are structural vectors of the characteristic plane $\tilde{\mathfrak{A}}_1'$ which, however, has a part (of $k - \chi$ dimensions) in common with the plane $\mathfrak{A}_1$ constructed earlier.

Let us now take into account the fact that

$$(A - \lambda)a_i' = a_{i-1}' \qquad (i = 1, 2, \cdots, \chi)$$

(since $\sigma_0 = \cdots = \sigma_{\chi-1} = 0$) and that the vectors

$$a_1', a_2', \cdots, a_\chi'$$

are consequently structural vectors of the characteristic plane $\mathfrak{A}_1'$ having no part in common with the characteristic plane $\tilde{\mathfrak{A}}_1'$,

just constructed.

Thus in case $\chi < k$ one can also construct, by utilizing the vectors $A^i a$ and $A^j a'$, two characteristic planes belonging to the one root λ_1 and together filling out a plane of dimension $(k - \chi) + k' + \chi = k + k'$, i.e., equal to the multiplicity with which the root λ_1 enters the polynomial $\phi(\lambda)\phi'(\lambda)$. By carrying through the indicated construction for each root of the polynomial $\phi'(\lambda)$, we shall have decomposed the entire plane $\mathfrak{A} + \mathfrak{A}'$ into characteristic planes, and with this we shall have solved the problem before us in the case when $\mathfrak{A} + \mathfrak{A}'$ is the whole n-dimensional space. The characteristic polynomial will in this case equal the polynomial[1] $\phi(\lambda)\phi'(\lambda)$.

- - - - - - - - - - - - - -

[1]The computation of the coefficients of the polynomial $\phi'(\lambda)$ reduces, obviously, to the determination of the coefficients of the resolution of the vector $A^{m'}a$ in terms of the linearly independent vectors (6). One may easily satisfy oneself that in case $m + m' = n$, the computation of the coefficients of the polynomial $\phi'(\lambda)$ may be carried out by means of the formulas

$$\alpha_i' = (a, Aa, \cdots, A^{m-1}a, a', Aa', \cdots, A^{m'-i-1}a', A^{m'-i+1}a', \cdots, A^{m'}a')\delta_i^{-1}$$
$$(i = 1, 2, \cdots, m') \quad ,$$

where

$$\delta_i = (-1)^i(a, Aa, \cdots, A^{m-1}a, a', \cdots, A^{m'-1}a') \quad .$$

These may thus be considered to be a generalization of the formulas given in the footnote to page 236 for the case that was omitted from consideration in the article by Acad. A. N. Krylov, [1].

§7. In conclusion we dwell on the case when the plane $\mathfrak{A} + \mathfrak{A}'$ does not exhaust the entire space. Having selected the arbitrary vector a", not lying in the plane $\mathfrak{A} + \mathfrak{A}'$, we construct the system of vectors

$$Aa'', A^2a'', \dots, A^{m''-1}a'' \quad ,$$

such that the vectors

(12) $a, Aa, \dots, A^{m-1}a, a', Aa', \dots, A^{m'-1}a', a'', Aa'', \dots, A^{m''-1}a''$

are linearly independent, but the equation

$$(A^{m''}a'' + \alpha_1''A^{m''-1}a'' + \dots + \alpha''_{m''-1}Aa'' + \alpha''_{m''}a'') + (\beta_1'A^{m'-1}a'$$
$$+ \dots + \beta'_{m'-1}Aa' + \beta'_{m'}a') + (\gamma_1A^{m-1}a + \dots + \gamma_{m-1}Aa + \gamma_m a) = 0$$

holds. The plane defined by the vectors (12) we shall denote by $\mathfrak{A} + \mathfrak{A}' + \mathfrak{A}''$.

Let us introduce into the discussion the polynomials

$$\phi''(\lambda) = \lambda^{m''} + \alpha_1''\lambda^{m''-1} + \dots + \alpha''_{m''-1}\lambda + \alpha''_{m''} \quad ,$$
$$\psi'(\lambda) = \beta_1'\lambda^{m'-1} + \dots + \beta'_{m'-1}\lambda + \beta'_{m'} \quad ,$$
$$\chi(\lambda) = \gamma_1\lambda^{m-1} + \dots + \gamma_{m-1}\lambda + \gamma_m \quad .$$

(Subsequent reasoning will show that if the region $\mathfrak{A} + \mathfrak{A}' + \mathfrak{A}''$ is the entire n-dimensional space, the characteristic equation of the affinor A has the following form:

$$\phi(\lambda)\phi'(\lambda)\phi''(\lambda) = 0 \quad .$$

The extension of this method of constructing the characteristic equation to the case when the region $\mathfrak{A} + \mathfrak{A}' + \mathfrak{A}''$ does not fill out the entire space is thus made quite obvious.)

Let λ_1 be a root of multiplicity k" ($\neq 0$) of the polynomial $\phi''(\lambda)$ and simultaneously a root of multiplicity k_2 of the polynomial $\phi'(\lambda)$ and of multiplicity k_1 of the polynomial $\phi(\lambda)$.

By means of the operators D and P, defined by formulas (3) and (7), and also of operator P', defined by the formula

$$P'\omega(\lambda) = \omega(\lambda) - \omega(\lambda_1)\frac{\psi'_{k_2}(\lambda_1)}{\phi'_{k_2}(\lambda_1)} ,$$

let us construct the system of vectors a''_o, a''_1, a''_2, $\cdots$, determined by the formula

$$a''_i = \phi''_i a'' + \psi'_i a' + \chi_i a \qquad (i = 1, 2, \cdots, k'') ,$$

where the polynomials ϕ''_i, ψ'_i, χ_i are defined by the following recurrence relation:

$$\phi''_{i+1} = D\phi''_i \ ; \qquad \phi''_o = \phi'' ,$$

(13) $$\psi'_{i+1} = DP'\psi'_i \ ; \qquad \psi'_o = \psi' ,$$

$$\chi_{i+1} = DP\left(\chi_i - \psi'_i(\lambda_1)\frac{\psi_{k_2}}{\phi'_{k_2}(\lambda_1)}\right) ; \qquad \chi_o = \chi .$$

Let us construct also the two systems of numbers

$$\sigma'_i = \psi'_i(\lambda_1)\cdot\frac{-1}{\phi'_{k_2}(\lambda_1)}$$

(13') $$\tau_i = (\chi_i(\lambda_1) - \sigma_i\cdot\psi_{k_2}(\lambda_1))\cdot\frac{-1}{\phi_{k_1}(\lambda_1)} \qquad (i = 0, 1, \cdots, k''-1) .$$

It can be shown that the following system of equations holds:

(14) $$(A - \lambda_1)a''_i = a''_{i-1} + \sigma'_{i-1}a'_{k_2} + \tau_{i-1}a_{k_1} \qquad (i = 1, 2, \cdots, k''; \ a''_o = 0).$$

Now let us construct the system of vectors c_i, defined by the recurrence relation

(15) $$(A - \lambda)c_i = c_{i-1} \ ; \quad c_{k''} = a''_{k''} \quad .$$

Employing (14), (11) and (5), we obtain[1]:

$$c_i = a''_i + \sigma'_i b_{\tilde{k}'} + \sigma'_{i-1} b_{\tilde{k}'-1} + \cdots + \sigma_{k''-1} a_{\tilde{k}'-(k''-i-1)}$$

$$+ \tau_i a_k + \tau_{i-1} a_{k-1} + \cdots + \tau_{k''-1} a_{k-(k''-i-1)}$$

and in particular,

$$c_1 = a''_1 + \sigma'_1 b_{\tilde{k}'} + \sigma'_2 b_{\tilde{k}'-1} + \cdots + \tau_1 a_k + \tau_2 a_{k-1} + \cdots$$

(here and henceforth it is assumed that $b_k = a_k = 0$ if $k \leqq 0$).

Therefore

(16) $$c_o = (\sigma'_o b_{\tilde{k}'} + \sigma'_1 b_{\tilde{k}-1} + \cdots) + (\tau_o a_{\tilde{k}} + \tau_1 a_{\tilde{k}-1} + \cdots) \quad .$$

a) If

$$b_{\tilde{k}'} = a'_{k'} = 0 \quad \text{and} \quad a_k = 0 \quad ,$$

or

$$\tilde{b}_{\tilde{k}} = 0 \ ; \quad \tau_o = \tau_1 = \cdots = \tau_{k''} = 0 \quad ,$$

or

$$a_k = 0 \ ; \quad \sigma'_o = \sigma'_1 = \cdots = \sigma'_{k''} = 0 \quad ,$$

or

$$\sigma'_o = \sigma'_1 = \cdots = \sigma'_{k''} = 0 \quad \text{and} \quad \tau_o = \tau_1 = \cdots = \tau_{k''} = 0 \quad ,$$

- - - - - - - - - - - - - -

[1] Here $\tilde{k}'$ is equal either to k' or to $k' + (k - \chi)$; in the latter case b_i denotes the vector that we have previously denoted by $b_{i-(k-\chi)}$.

then $c_o = 0$ and the vectors

$$c_1;\ c_2;\ \cdots;\ c_{k''} \qquad (c_i = a_i'')$$

are structural vectors of the characteristic plane α_1'', which belongs to the characteristic number λ_1 and has no common part with the characteristic planes constructed earlier.

b) If, furthermore, $b_{\tilde{k}} = 0$ (or $\sigma_o' = \cdots = \sigma_{k''-1}' = 0$), but $a_k \neq 0$ and $\tau_o = \tau_1 = \cdots = \tau_{s-1} = 0; \quad \tau_s \neq 0\ ; \quad 0 \leqq s < k''$,

we then obtain (in conformity with (16))

$$c_o = \tau_s a_{k-s} + \tau_{s+1} a_{k-s-1} + \cdots \quad .$$

If $k - s \leqq 0$, then $c_o = 0$, and accordingly the vectors

$$c_1,\ c_2,\ \cdots,\ c_{k''} \qquad (c_1 = a_1'' + \tau_s a_{k-s+1})$$

are structural vectors of the characteristic plane α_1'' which has no common part with the planes previously constructed.

If $k - s > 0$, however, then $b_o \neq 0$ and, continuing the construction prescribed by the recurrence formula (15), we construct the vectors

$$c_{-1},\ c_{-2},\ \cdots,\ c_{-(k-s)+1} = \tau_s a_1 \quad ;$$

the (linearly independent) vectors

$$c_{-(k-s)+1},\ c_{-(k-s)+2},\ \cdots,\ c_o,\ c_1,\ \cdots,\ c_{k''}$$

are structural vectors of the characteristic plane $\tilde{\alpha}_k''$ having $k'' + (k - s)$ dimensions, which has, however, a common part (of $k - s$ dimensions) with the previously constructed plane α_1.

Let us now take into account the fact that in conformity with (14),

$$(A - \lambda)a_i'' = 0 \qquad (i = 1, 2, \cdots, s)$$

and accordingly the vectors

$$a_1'', a_2'', \cdots, a_s''$$

are structural vectors of the characteristic plane α_1' which does not have a common part with the characteristic planes constructed earlier. Thus in the case under consideration it is likewise possible to construct, by using the vectors $A^i a$, $A^j a'$, $A^{\gamma} a''$, three characteristic planes belonging to the same characteristic number λ_1 and together filling out a plane of dimension $[k'' + (k' - s)] + s + k' = k + k' + k''$, i.e., of the multiplicity with which the root λ_1 figures in the polynomial $\phi(\lambda)\phi'(\lambda)\phi''(\lambda)$.

c) We shall not tarry over a consideration of the case when

$$a_k = 0 \ , \quad \text{or} \quad \tau_0 = \tau_1 = \cdots = \tau_{k''-1} = 0 \quad ,$$

but

$$a'_{k'} = b_{\tilde{k}} \neq 0 \ ; \quad \sigma'_0 = \sigma'_1 = \cdots = \sigma'_{\chi'-1} = 0 \ ; \quad \sigma'_{\chi'} \neq 0 \ ; \quad 0 \leqq \chi' < k'' \ .$$

Let us pass on to the case when

$$a_k \neq 0 \ ; \quad \tau_0 = \tau_1 = \cdots = \tau_{s-1} = 0 \ ; \quad \tau_s \neq 0 \ ; \quad 0 \leqq s < k'' \ ,$$

$$b_{\tilde{k}} \neq 0 \ ; \quad \sigma'_0 = \sigma'_1 = \cdots = \sigma_{\chi'-1} = 0 \ ; \quad \sigma_{\chi'} \neq 0 \ ; \quad 0 \leqq 1' < k'' \ .$$

In accordance with (16) we obtain:

$$c_0 = (\sigma_{\chi'} b_{\tilde{k}-\chi'} + \sigma_{\chi'+1} b_{\tilde{k}-\chi'-1} + \cdots) + (\tau_s a_{k-s} + \tau_{s+1} a_{k-s+1} + \cdots) \ .$$

If

$$\tilde{k} - \chi' \leqq 0 \quad \text{and} \quad k - s \leqq 0 \quad ,$$

then $c_0 = 0$ and accordingly the vectors

$$c_1, c_2, c_3, \cdots, c_{k''} = a_{k''}$$

are structural vectors of the characteristic plane.

Omitting consideration of the case when either $\tilde{k} - \chi' \leqq 0$ or $k - s \leqq 0$, let us turn to the case when

$$\tilde{k} - \chi' > 0 \quad \text{and} \quad k - s > 0 \quad .$$

Here let

$$\tilde{k} - \chi' \geqq k - s \quad .$$

Continuing the construction prescribed by the recurrence formula (15), let us construct the linearly independent vectors

$$c_{-1}, c_{-2}, \cdots, c_{-(\tilde{k}-\chi')+1} = \sigma'_{\chi'} a'_1 \quad ;$$

the vectors

$$c_{-(\tilde{k}-\chi')+1}, c_{-(\tilde{k}-\chi')+2}, \cdots, c_o, c_1, \cdots, c_{k''}$$

are structural vectors of the characteristic plane $\mathfrak{A}''_1$ (which has $k'' + (\tilde{k} - \chi')$ dimensions) which has, however, a common part of $\tilde{k} - \chi'$ dimensions with the characteristic plane constructed earlier, whose structural vectors are headed by the vector a'_1.

c_1) If with this we have $\chi' \leqq s$, then in accordance with (14),

$$(A - \lambda)a''_i = a''_{i-1} \qquad (i = 1, 2, \cdots, \chi') \quad ,$$

and accordingly the vectors

$$a''_1, a''_2, \cdots, a''_{\chi}$$

are structural vectors of a characteristic plane that has no common part with the characteristic planes constructed previously. Thus in the case in hand one can construct three characteristic

planes belonging to the same characteristic number and together filling out the plane whose dimensions are equal to

$$k'' + (\tilde{k}' - \chi') + \chi' + \tilde{k} = k'' + \tilde{k}' + \tilde{k} = k'' + k' + k \quad .$$

c_2) If, however, $\chi' > s$, then in accordance with (14) the equations

$$(A - \lambda)a''_i = a''_{i-1} \qquad (i = 1, 2, \cdots, s) \quad ,$$

$$(A - \lambda)a''_i = a''_{i-1} + \tau_{i-1}a_k \qquad (i = s + 1, \cdots, \chi')$$

hold. In this case let us construct a system of vectors a_i, defined by the recurrence formula

(17) $$(A - \lambda)d_i = d_{i-1} \; ; \quad d_{\chi'} = a''_{\chi''} \quad .$$

Then

$$d_i = a''_i + \tau_i a_k + \tau_{i+1}a_{k-1} + \cdots$$

and, in particular,

$$d_1 = a''_1 + \tau_1 a_k + \tau_2 a_{k-1} + \cdots \quad .$$

Therefore

$$d_o = \tau_o a_k + \tau_1 a_{k-1} + \cdots \quad ,$$

and accordingly

$$d_o = \tau_s a_{\tilde{k}-s} + \tau_{s+1}a_{\tilde{k}-s+1} + \cdots \quad .$$

Since $\tilde{k} - s > 0$, $d_o \neq 0$, and we construct, by means of the recurrence formula (17), the linearly independent vectors

$$d_{-1};\; d_{-2};\; \cdots;\; d_{-(\tilde{k}-s)+1} = \sigma_{\chi'}a_1 \quad .$$

The vectors

$$d_{-(\tilde{k}-s)+1}, \cdots, d_{-1}, d_o, d_1, \cdots, d_{\chi'}$$

are structural vectors of the characteristic plane (having a common part of $\tilde{k} - s$ dimensions with the plane constructed previously that is headed by the vector a_1).

Let us now take into consideration the fact that

$$(A - \lambda)a_i'' = a_{i-1}'' \qquad (i = 1, 2, \cdots, s) \quad .$$

Therefore the vectors

$$a_1'', a_2'', \cdots, a_s''$$

are structural vectors of the characteristic plane. Thus in the case indicated one can construct three characteristic planes, which together fill out a plane of dimension

$$[k'' + (\tilde{k}' - \chi')] + [\chi' - (\tilde{k} - s)] + s = k'' + \tilde{k}' + \tilde{k} = k'' + k' + k \ ,$$

i.e., of the multiplicity with which the root λ_1 figures in the polynomial $\phi(\lambda)\phi'(\lambda)\phi''(\lambda)$.

Having carried through the indicated constructions for each root of the polynomial $\phi''(\lambda)$, we shall have decomposed the whole plane $\alpha + \alpha' + \alpha''$ into characteristic planes and shall with this have solved the problem in hand in the case where $\alpha + \alpha' + \alpha''$ constitutes the whole n-dimensional space. The characteristic polynomial will in this case obviously be equal to the polynomial $\phi(\lambda)\ \phi'(\lambda)\phi''(\lambda)$.

We shall not dwell on a discussion of that case for which the plane $\alpha + \alpha' + \alpha''$ is only part of the n-dimensional space; the reasoning set forth above indicates the course that would have to be followed for a complete separation of the n-dimensional space into structural planes.

§8. a) The case $m < n$, discussed in §1, is, theoretically speaking, exceptional. It can occur either with an exceptional structure of the affinor A (where to some characteristic number of A there correspond different characteristic planes), or with an exceptional choice of the initial vector $\underline{a}$ (when it is taken from some characteristic plane of the affinor A). One must suppose this to be the explanation of the negligent attitude toward consideration of this case that characterizes the literature of the problem.

It is therefore essential to show that "in practice", i.e., in computations conducted to a limited degree of accuracy, this case will be <u>general</u> when n is sufficiently large.

In order to elucidate this, let us denote by $\lambda_1, \cdots, \lambda_{n_1}$ the group of characteristic numbers of the operator A that are largest in modulus, and assume that the remaining characteristic numbers $\lambda_{n_1+1}, \lambda_{n_1+2}, \cdots, \lambda_{n_1+n_2}$ "have the order of smallness k with respect to the number λ_1", i.e., for the adopted degree of computational accuracy, the fraction $\left(\frac{\lambda_{n_1+i}}{\lambda_1}\right)^k$ may be disregarded in comparison with unity.

Assuming for simplicity of exposition that the characteristic numbers λ_i are simple and distinct, and denoting by p_i the characteristic vectors corresponding to them, we obtain

$$A^s a = \sum_{\mu=1}^{n} a^\mu \lambda_\mu^s p_\mu \quad ,$$

if

$$a = \sum_{\mu=1}^{n} a^\mu p_\mu \quad .$$

(The coordinates α^i of the vector $\underline{a}$ are generally speaking numbers of the same order.)

On the strength of the assumption that was made,

$$1 + \left(\frac{\lambda_{n_1+1}}{\lambda_1}\right)^k \approx 1 \quad ,$$

we arrive at the conclusion that the vector

$$b \equiv A^k a \approx \sum_1^{n_1} a^\mu \lambda_\mu^k p_\mu \quad ,$$

i.e., "in practice" it lies in the plane formed by the "first"[1] characteristic vectors $p_1, p_2, \cdots, p_{n_1}$, and that the vectors $Ab, A^2b, \cdots, A^{n_1-1}b$ lie in this same plane, and consequently the vectors

$$b, Ab, \cdots, A^{n_1-1}, A^{n_1}b$$

are "in practice" connected by a relation of linear dependence:

$$\omega(A)b \equiv A^{n_1}b + \alpha_1 A^{n_1-1}b + \cdots + \alpha_{n-1}Ab + \alpha_n b = 0 \quad .$$

Thus in the case when

$$k + n_1 < n$$

(and this is just what will be true, generally speaking, when n is sufficiently large), the vectors

$$a, Aa, \cdots, A^{k-1}a, A^k a, \cdots, A^{k+n_1}a$$

are linearly dependent and accordingly the method indicated in §1 will lead to the determination only of the "first" characteristic numbers and the characteristic vectors corresponding to them.

It is easily realized that in case $n_1 = 1$, we arrive at the well-known "method of iteration" for the determination of the largest characteristic number and the characteristic vector corresponding to it.

- - - - - - - - - - - - - -

[1]Editor's note: The author uses the word for "oldest".

It is essential to note that the recommended practice for this method,[1] under which one verifies at each succeeding stage only the collinearity of the two successive vectors $A^i a$ and $A^{i-1}a$, artificially reduces the general method proposed by us to the case $n_1 = 1$.

A more expeditious method in practice is thus the following computational scheme, which derives from the reasoning set forth above, and offers the possibility of discovering the whole plane of n_1 dimensions formed by the first characteristic vectors. _Having constructed the critical vector $A^i a$, let us determine the possibility of resolving it in terms of the vectors constructed earlier._

$$A^{i-1}a,\ A^{i-2}a,\ \ldots,\ A^2a,\ A^1a,\ a \quad .$$

Taking into consideration along with this the practical likelihood that the vector $A^i a$ will turn out to be a linear combination of just some of the vectors immediately preceding it in this sequence, we should employ with this aim in view the "second scheme of linear analysis".[2]

b) We shall now indicate a possible course of the determination of the rest of the characteristic vectors and characteristic numbers.

We note to begin with that the course suggested in §3 will not lead to the goal in the case in hand, since -- in conformity with what has been stated in §3 -- having again begun the process

- - - - - - - - - - - - - -

[1]See [6].

[2]See below, §9.

of iterations, proceeding from a vector a' which does not lie in the plane $p_1, p_2, \cdots, p_{n_1}$ already constructed, we in practice arrive again at the plane $p_1, p_2, \cdots, p_{n_1}$, since the vector $A^k a'$ will "practically" lie in this very plane.

The process suggested in §6 will also not give the possibility of finding the new characteristic vectors. Indeed, with the construction of the n_1+k vectors

$$b,\ Ab,\ \cdots,\ A^{n_1-1}b,\ a',\ Aa',\ \cdots,\ A^{k-1}a'$$

we will not yet have obtained a basis, since $n_1 + k < n$; at the same time the next vector constructed will "practically" lie in the plane of the vectors already constructed, $p_1, \cdots, p_n$.

The position would, however, have been altered had the vector a' been taken from the plane of the "last"[1] characteristic vectors $p_{n_1+1}, p_{n_1+2}, \cdots, p_{n_1+n_2}$. In this case the scheme indicated above would have led us ("theoretically") to the determination of a certain plane (the "second"[2]), formed by the vectors

$$p_{n_1+1},\ p_{n_1+2},\ \cdots,\ p_{n_1+k'} \qquad (k' < n_1) \quad .$$

The vector $\omega(A)a$ may be taken as such a vector a'. Indeed, having taken into consideration that

$$\omega(A)p_i = 0 \qquad (i = 1, 2, \cdots, n_1) \quad ,$$

we obtain the result

$$\omega(A)a = \alpha^{n_1+1}\omega(\lambda_{n_1+1})p_{n_1+1} + \cdots + \alpha^n\omega(\lambda_n)p_n \quad .$$

- - - - - - - - - - - - - -

[1]Editor's note: The author uses the word for "youngest".
[2]Editor's note: The author uses the word for "second oldest".

It should, however, be borne in mind that "practically" the vector $A^k a'$ will also contain components with respect to the vectors $p_1, p_2, \cdots, p_{n_1}$; however, they will be small in comparison with its components with respect to the vectors $p_{n_1+1}, \cdots, p_{n_1+n_2}$. One should therefore expect that despite the fact that the vector $A^k a'$ will have a component in the plane $p_1, \cdots, p_{n_1}$, it will not be so large that one can neglect, in comparison with it, the component of the vector $A^k a'$ in the plane $p_{n_1+1}, p_{n_1+2}, \cdots, p_n$.

Continuing the computation of the vectors

$$A^{k+1}a', A^{k+2}a', A^{k+3}a', \cdots$$

we will of course obtain vectors in which the component in the plane $p_1, p_2, \cdots, p_{n_1}$ increases in comparison with the component in the plane $p_{n_1+1}, \cdots, p_{n_1+n_2}$. One may anticipate, however, that in some cases the vector

$$A^{n-n_1}a' \equiv A^{k+(n-n_1-k)}a'$$

will no longer lie in the plane $p_1, \cdots, p_n$, and therefore the vectors

$$b_1, Ab_1, \cdots, A^{n_1-1}b_1, a_1', Aa_1', \cdots, A^k a', A^{k+1}a', \cdots, A^{n-n_1-1}a'$$

will form a basis of the space. Resolving the vector $A^{n-n_1}a'$ in terms of the vectors of this basis, we arrive at the equation

$$\phi'(A)a' + \psi(A)b = 0 \quad ,$$

by means of which we will find the characteristic vectors $p_{n_1}, \cdots, p_n$ as has been shown in §6.

§9. The method given above for the determination of characteristic numbers and characteristic vectors leads to the need for a practical solution of the following problem (the problem of "linear analysis").

<u>In the sequence of vectors</u>

$$a_1, a_2, a_3, \dots ,$$

each of which is given in some coordinate system $\chi_1, \chi_2, \dots , \chi_n$ by its coordinates, $a_i = \alpha^{n}_{i\mu}\chi_{\mu}$: <u>to find the vector of least index that is linearly expressible in terms of the preceding vectors, and to compute the coefficients of this resolution.</u>

We shall indicate two schemes for solving this problem. The first of them is convenient for the case when n (the number of dimensions of the space) is small. The second -- somewhat inferior in point of computational convenience -- substantially reduces the number of operations in cases for which n is large.[1]

<u>First scheme</u>

Given the vectors

---

[1] Editor's note: Let M denote the matrix whose rows are the vectors $a_1, a_2, \dots$. Then the author's first scheme is ordinary Gaussian elimination on the rows of M to triangularize M. The second scheme is a modified Gaussian elimination on the columns of A. In both schemes the author deals carefully with zeros. He fails, however, to note the practical efficiency achieved in the compact arrangements of elimination introduced by many computers and described, for example, in P.S. Dwyer's <u>Linear Computations</u> (John Wiley, 1951).

$$a_1, a_2, a_3, \cdots \qquad (a_k \equiv a_k^1) \quad ,$$

which we shall call the vectors of the first series, let us construct the vectors

$$a_1^2, a_2^2, a_3^2, \cdots \quad , \qquad a_i^2 = \alpha_i^{2\mu}\chi_\mu$$

of a second series, by the formulas

$$a_m^2 = a_{m+1}^1 - a_1^1\lambda_m^1 \quad ;$$

here

$$\lambda_m^1 = \alpha_{m+1}^{1k_1} : \alpha_1^{1k_1} \; ,$$

where $\alpha_1^{1k_1}$ is the first (i.e., with the least index) non-vanishing coordinate of the vector a_1. The vectors of the second series obviously lie in an $(n-1)$-dimensional coordinate plane defined by the coordinate vectors

$$\chi_1, \chi_2, \cdots, \chi_{k_1-1}, \chi_{k_1+1}, \cdots, \chi_n \quad .$$

If the first vector of the second series, a_1^2, is not equal to zero, then construct the vectors of a third series:

$$a_1^3, a_2^3, a_3^3, \cdots$$

by the formulas

$$a_m^3 = a_{m+1}^2 - a_1^2\lambda_m^2 \quad ;$$

here

$$\lambda_m^2 = \alpha_{m+1}^{2k_2} : \alpha_1^{2k_2} \; .$$

where $\alpha_1^{2k_2}$ is the first non-vanishing coordinate of the vector a_1^2. If $a_1^3 \neq 0$, then construct analogously the vectors $a_1^4, a_2^4, \cdots$ of a fourth series, etc., thus obtaining successively the elements of a triangular vector table

$$\text{(I)}\qquad \left.\begin{array}{ccccccc} a_1^1 & a_2^1 & a_3^1 & \cdots & & a_{p-1}^1 & a_p^1 \\ & a_1^2 & a_2^2 & & & a_{p-2}^2 & a_{p-1}^2 \\ & & a_1^3 & & & a_{p-3}^3 & a_{p-2}^3 \\ & & & \ddots & & & \\ & & & & a_1^s \cdots & a_{p-s}^s & a_{p-s+1}^s , \end{array}\right\}$$

and also the triangular table of numbers

$$\text{(I')}\qquad \left.\begin{array}{ccccc} \lambda_1^1 & \lambda_2^1 & \lambda_3^1 & \cdots & \lambda_{p-1}^1 \\ & \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_{p-2}^2 \\ & & \lambda_1^3 & \cdots & \lambda_{p-3}^3 \\ & & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot \\ & & \lambda_1^s & \cdots & \lambda_{p-s}^s , \end{array}\right\}$$

defined by the recurrence formulas

$$a_m^{i+1} = a_{m+1}^i - a_1^i \lambda_m^i \quad ;$$

$$\lambda_m^i = \alpha_{m+1}^{ik_i} : \alpha_1^{ik_i} \quad ,$$

where $\alpha_1^{ik_i}$ is the first non-vanishing coordinate of the vector a_1^i.

If the first vector of the (s + 1)-th series a_1^{s+1} turns out to be equal to zero, then, as one may easily verify,

$$a_{s+1} = \xi^1 a_1 + \xi^2 a_2 + \cdots + \xi^s a_s \quad ,$$

where the coefficients ξ^i of the resolution are determined from the triangular system of linear equations

$$\xi^p + \xi^{p+1}\lambda_1^p + \xi^{p+2}\lambda_2^p + \cdots + \xi^{s-1}\lambda_{s-p-1}^p + \xi^s\lambda_{s-p}^p = \lambda_{s-p+1}^p$$

$$(p = 1, 2, \cdots, s) \; ; \; \lambda_o^i = 1 \quad ,$$

which is written in extenso thus:

$$\text{(II)} \quad \left.\begin{array}{l} \xi^1 + \xi^2\lambda_1^1 + \xi^3\lambda_2^1 + \xi^4\lambda_3^1 + \cdots + \xi^{s-1}\lambda_{s-2}^1 + \xi^s\lambda_{s-1}^1 = \lambda_s^1 \\ \qquad \xi^2 + \xi^3\lambda_1^2 + \xi^4\lambda_2^2 + \cdots + \xi^{s-1}\lambda_{s-3}^2 + \xi^s\lambda_{s-2}^2 = \lambda_{s-1}^2 \\ \qquad\qquad \xi^3 + \xi^4\lambda_1^3 + \cdots + \xi^{s-1}\lambda_{s-4}^s + \xi^s\lambda_{s-3}^3 = \lambda_{s-2}^3 \\ \qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad \xi^{s-1} + \xi^s\lambda_1^{s-1} = \lambda_2^{s-1} \\ \qquad\qquad\qquad\qquad \xi^s = \lambda_1^s \; . \end{array}\right\}$$

It is useful to carry through the indicated computations in such a sequence that the triangular table (I) of the vectors a_i^k develops in sequential column construction; speaking more exactly, after constructing the vector a_1^2 (which completes the second column) the vectors a_2^2, a_1^3 are constructed in succession (completing the third column), and then the vectors a_3^2, a_2^3, a_1^4, etc.

If not one vector equal to zero appears in the first s columns, this will imply that the vectors $a_1, a_2, a_3, \cdots, a_s$ are linearly independent.

If, furthermore, the vector a_{s-k+2}^k, in the $(s+1)$-th column, turns out to be zero (and the vectors standing above it are different from zero), this will then imply that

$$a_{s+1} = \xi^1 a_1 + \cdots + \xi^k a_k$$

(since in this case all vectors of the $(s+1)$-th column standing

beneath the vector a^k_{s-k+2} will equal zero, and consequently so will $a^{s+1}_1 = 0$).

The coefficients ξ^i of the resolution are in this case solutions of a foreshortened system of equations that is obtained from the system (II) of equations, if we put in it

$$\lambda^{k+1}_{s-k} = \lambda^{k+2}_{s-k-1} = \cdots = \lambda^s_1 = 0 \quad .$$

Observation. The suggested scheme obviously offers the possibility of investigating and solving the system of n equations in p unknowns:

$$\xi^1\alpha^k_1 + \xi^2\alpha^k_2 + \cdots + \xi^p\alpha^k_p = \alpha^k_{p+1} \qquad (k = 1, 2, \cdots, n) \quad ,$$

since it is equivalent to one vector equation

$$\xi^1 a_1 + \xi^2 a_2 + \cdots + \xi^p a_p = a_{p+1} \; ; \qquad a_i = a^{\mu}_i \gamma_{\mu} \quad .$$

If $p = n$, and if it is known beforehand that the vectors $a_1, a_2, \cdots, a_n$ are linearly independent, it is then convenient (from a computational point of view) to conduct the construction of the triangular vector table so that the rows develop successively in it (i.e., first compute the vectors of the second series, then of the third, and so forth).

It is readily calculated that for the construction of tables (I) and (I'), one must carry out

$$\tfrac{1}{2}(n-1)p(p-1) \begin{array}{l}\text{additions}\\ \text{subtractions}\end{array}\Bigg\} \quad \text{and} \quad \tfrac{1}{2}np(p-1) \begin{array}{l}\text{multiplications}\\ \text{divisions}\end{array}\Bigg\} \; .$$

For computing the coefficients ξ^i of the triangular system of equations (II) one must still carry out

$$\frac{1}{2}s(s-1) \left.\begin{matrix}\text{additions}\\ \text{subtractions}\end{matrix}\right\} \quad \text{and} \quad \frac{1}{2}s(s-1) \left.\begin{matrix}\text{multiplications}\\ \text{divisions}\end{matrix}\right\}.$$

Thus for the solution of the system of n independent linear equations in n unknowns by the scheme suggested here, one must carry through $(n = p - s)$ in all

$$\frac{1}{2}(n-1)n^2 \left.\begin{matrix}\text{additions}\\ \text{subtractions}\end{matrix}\right\} \quad \text{and} \quad \frac{1}{2}(n-1)n(n+1) \left.\begin{matrix}\text{multiplications}\\ \text{divisions}\end{matrix}\right\}.$$ [1]

The execution of the operations according to the indicated scheme is much facilitated if specially constructed cut-out templates be utilized.

Second scheme.

Given the vectors $a_1^1, a_2^1, a_3^1, \dots$ of the "first series", let us construct the vectors $a_1^2, a_2^2, a_3^2, \dots$ of a second series by the formulas

$$a_m^2 = a_{m+1}^1 - a_m^1 \lambda_m^1 \;;$$

here

$$\lambda_m^1 = \alpha_{m+1}^1 : a_m^1 \;.$$

The vectors of the second series obviously lie in the plane defined by the vectors $\chi_2, \chi_3, \dots, \chi_n$. Construct, further, the vectors of

- - - - - - - - - - - - -

[1]C. Runge's scheme (Praxis der Gleichungen) requires the same number of operations; it consists of the successive elimination of the unknown ξ^1 from all the equations commencing with the second, then ξ^2 from all equations commencing with the third, etc. We note, however, that Runge's scheme does not lead to results in the case where there are linearly dependent or incompatible equations in the system.

a third series: a_i^3 $(i = 1, 2, 3, \cdots)$ by the formulas:

$$a_m^3 = a_{m+1}^2 - a_m^2 \lambda_m^2 \quad ;$$

here

$$\lambda_m^2 = \alpha_{m+1}^{22} : \alpha_m^{22} \quad .$$

These vectors obviously lie in the plane spanned by the vectors $\chi_3, \chi_4, \cdots, \chi_n$. We construct the vectors a_i^4 of a fourth series, those of a fifth series, and so on, obtaining the elements of a triangular vector table

$$\text{(III)} \qquad \left. \begin{array}{cccc} a_1^1 & a_2^1 & a_3^1 & a_4^1 \cdots \\ & a_1^2 & a_2^2 & a_3^2 \cdots \\ & & a_1^3 & a_2^3 \cdots \\ & & & \cdots \end{array} \right\}$$

and also the triangular table of numbers

$$\text{(III')} \qquad \left. \begin{array}{cccc} \lambda_1^1 & \lambda_2^1 & \lambda_3^1 & \lambda_4^1 \cdots \\ & \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \cdots \\ & & \lambda_1^3 & \lambda_2^3 \cdots \\ & & & \cdots \end{array} \right\}$$

defined by the recurrence formulas

$$a_m^{k+1} = a_{m+1}^k - a_m^k \lambda_m^k \quad ;$$

$$\lambda_m^k = \alpha_{m+1}^{kk} : \alpha_m^{kk}$$

(here obviously $a_p^k = a_p^{k\mu} \chi_\mu$).

The indicated construction may conveniently be carried through

in such an order that the triangular table of vectors (III) develops by a sequential column construction (just as was suggested for the first scheme).

If a vector equal to zero does not appear in the first s columns, this will imply that the vectors $a_1, a_2, \cdots, a_s$ (of the first series) are linearly independent.

Furthermore if the vector a^{k+1}_{s-k+1} of the (s + 1)-th column turns out to be equal to zero (but the vectors standing above it are different from zero), this will imply that

$$a_{s+1} = \xi^s a_s + \xi^{s-1} a_{s-1} + \cdots + \xi^{s-k+1} a_{s-k+1} \quad ;$$

(if, in particular, only $a^{s+1}_1 = 0$, then $a_{s+1} = \xi^s a_s + \xi^{s-1} a_{s-1} + \cdots + \xi^1 a_1$).

For the determination of the coefficients ξ^i of the resolution it must be borne in mind that the equation for a^{k+1}_{s-k+1} has for its consequence, in view of the law of construction of the vector table (III), the sequence of equations

$$a^k_{s-k+2} = a^k_{s-k+1} \cdot \overset{k}{\xi}{}^{s-k+1} \; ; \quad \overset{k}{\xi}{}^{s-k+1} = \lambda^k_{s-k+1} \quad ,$$

$$a^{k-1}_{s-k+3} = a^{k-1}_{s-k+2} \cdot \overset{k-1}{\xi}{}^{s-k+2} + a^{k-1}_{s-k+1} \overset{k-1}{\xi}{}^{s-k+1} \quad ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a^m_{s-m+2} = a^m_{s-m+1} \cdot \overset{m}{\xi}{}^{s-m+1} + a^m_{s-m} \overset{m}{\xi}{}^{s-m} + \cdots + a^m_{s-k+1} \overset{m}{\xi}{}^{s-k+1} \quad ,$$

in which the coefficients $\overset{s}{\xi}{}^i$ are to be determined by means of the following recurrence formulas:

$$\overset{m}{\xi}{}^i = \overset{m+1}{\xi}{}^{i-1} - \overset{m+1}{\xi}{}^i \lambda^m_i \qquad \begin{matrix} m = k-1, k-2, \cdots, 2, 1 \\ i = s-k+1, s-k+2, \cdots, s-m+1 \end{matrix} \quad ,$$

in which we have put

$$\overset{m+1}{\xi}{}^{s-k} = 0 \; ; \qquad \overset{m+1}{\xi}{}^{s-m+k} = -1 \quad .$$

Computationally, the determination of the coefficients ξ^i ($= \overset{1}{\xi}{}^i$) from this system is less convenient than the solution of the triangular system of equation (II) (which appears in the first scheme); however the use of specially constructed templates will reduce this discrepancy to a minimum.

The superiority of the second scheme will manifest itself, however, in cases where in resolving the vector a_{s+1} in terms of the vectors $a_1, \cdots, a_s$ several of the first vectors $a_1, a_2, \cdots, a_k$ do not appear. This circumstance, which will obtain, as a rule, with the determination of the characteristic numbers and vectors of a matrix of high order (for the case of a large n; see §8), will be revealed automatically, in a sense, during the utilization of the second scheme (by the vanishing of the vectors $\overset{k+1}{a}_{s-k+1}$) -- more exactly speaking, where one does not need to compute the coefficients of the vectors $a_1, a_2, \cdots, a_k$ in the resolution.

The second scheme also offers the possibility of beginning the construction of the triangular vector table (III) -- in cases when it is possible to anticipate the absence, in the resolution of a_{s+1}, of components of the vectors $a_1, \cdots, a_k$ -- not with the vector a_1 but with the vector a_k: if the process of the computation suggests the wisdom of using the vectors $a_{k-1}, a_{k-2}, \cdots$, as well, a "development leftwards" of table (III) is possible only with the use of the second scheme.

BIBLIOGRAPHY

(All in Russian)

[1] A. N. Krylov, "The numerical solution of an equation with which the frequency of the vibrations of material systems is determined in technological problems", Akad. Nauk. SSSR, Izvestiia, Physico-mathematical series (1931), pp. 491-539.

[2] A. M. Danilevsky, "The numerical solution of the characteristic equation", Matem. Sbornik, vol. 2 (44), (1937), pp. 169-171.

[3] A. M. Lopshits, "The characteristic equation of lowest degree for an affinor, and its application to the integration of differential equations", Same Trudy,* vol. 2-3 (1935), p. 11.

[4] P. A. Shirokov, Tensor analysis. Part 1. The algebra of tensors, GTTI (1934).

[5] N. N. Luzin, "Concerning the method of A. N. Krylov for composing the characteristic equation", Akad. Nauk. SSSR, Izvestiia, Physico-mathematical series (1931), pp. 903-958.

[6] K. A. Semendiaev, "The determination of the latent roots and the invariant manifolds of a matrix by means of iterations", Priklad. matem. i mekh, vol. 3 (1943), pp. 193-221. [Available in English: National Bureau of Standards Report No. 1402, 1952.]

*Cited on page 1 above.

II

THE APPLICATION OF POLYNOMIALS OF BEST APPROXIMATION TO THE IMPROVEMENT OF THE CONVERGENCE OF ITERATIVE PROCESSES

by

M. K. Gavurin[1]

The purpose of this note is to propose some methods that will permit the acceleration of iterative processes in solving systems of linear algebraic equations and determining the proper numbers and vectors. We have, it should be said, limited ourselves to the algebraic case only for simplicity of exposition; it is perfectly evident that the methods proposed are applicable to analogous problems in the case of linear operators in an arbitrary Hilbert space.

1. Finding the proper numbers and vectors.

Let us consider a matrix of the n-th order, $A = (a_{ij})$, and assume that all its elementary divisors are linear and its proper numbers real.

Numbering the proper numbers of the matrix A in order of diminishing absolute value,

$$|\lambda_1| \geqq |\lambda_2| \geqq \dots \geqq |\lambda_n| \quad ,$$

let us in addition assume (which is no longer very essential) that

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \quad .$$

The ordinary iterative methods for finding the dominant proper

---

[1] Uspekhi Matematicheskhikh Nauk, vol. 5, no. 3 (1950), pp. 156-160.

number λ_1 consist of two stages:

1) The construction of a finite sequence $f_o, f_1, \cdots, f_p, f_{p+1}$ of numbers of the form

$$f_k = \sum_{i=1}^{n} \alpha_i \lambda_i^k \tag{1}$$

where α_i are certain constants. The sequence of traces of the matrices A^k can serve as such a sequence, or -- for an arbitrary vector Y_o, $Y_k = AY_{k-1} = A^k Y_o$ $(k = 1, 2, \cdots, p+1)$ -- the sequence of the i-th components of the vectors Y_k (i arbitrary), or lastly, in the case of a symmetric A, the sequence $f_{2k} = (Y_k, Y_k)$, $f_{2k+1} = (Y_k, Y_{k+1})$.

2) The approximate determination of λ_1, based on the fact that for $\alpha_1 \neq 0$ and k sufficiently large the first term will predominate in sum (1). Ordinarily one adopts

$$\lambda_1 \approx \phi_{p+1} = \frac{f_{p+1}}{f_p} .$$

As is easily seen, the error of this equality is estimable by the quantity

$$M_p = \frac{M\gamma^p(1 + \gamma)}{1 - M\gamma^p} .$$

where

$$M = \frac{1}{|\alpha_1|} \sum_{i=2}^{p} |\alpha_i| \ ; \qquad \gamma = \left|\frac{\lambda_2}{\lambda_1}\right| .$$

A. C. Aitken[1] has proposed the more exact equality

- - - - - - - - - - - - - - -

[1]A. C. Aitken, "Studies in practical mathematics. The evolution of the latent roots and the latent vectors of a matrix," Proc. Roy. Soc. of Edinburgh, vol. 37 (1937), pp. 269-304.

$$\lambda_1 \approx \frac{\phi_{p+1}\phi_{p-1} - \phi_p^2}{\phi_{p+1} - 2\phi_p + \phi_{p-1}} ,$$

the error of which is of the order $\gamma^{2p} + \left|\frac{\lambda_3}{\lambda_1}\right|^p$.

Knowing λ_1 and taking into consideration that in the difference $f_{k+1} - \lambda_1 f_k$ the term $\alpha_2\lambda_2^k(\lambda_2 - \lambda_1)$ predominates, λ_2, for example, may be found from the relation

$$\lambda_2 \approx \frac{f_{k+1} - \lambda_1 f_k}{f_k - \lambda_1 f_{k-1}} .$$

Here, in practice, one should not take k too large lest the inexactitude in the determination of λ_1 distort the result.

We have set ourselves the task of finding a linear combination

$$\theta_p = \sum_{k=o}^{p} c_k f_k = \sum_{i=1}^{n} \alpha_i \sum_{k=o}^{p} c_k \lambda_i^k = \sum_{i=1}^{n} \alpha_i P_p(\lambda_i) ,$$

such that the predominance of the first term in the right member will be as great as possible. We shall consider that λ_2 is known to us, and that nothing more of the distribution of $\lambda_3, \cdots, \lambda_n$ is known. We therefore arrive at the problem of minimizing the ratio

$$(2) \qquad \left|\frac{1}{P_p(\lambda_1)}\right| \cdot \max_{-|\lambda_2| \leq \lambda \leq |\lambda_2|} |P_p(\lambda)| .$$

It is easily seen that this ratio will attain its least value if

$$P_p(\lambda) = \sum_{k=o}^{p} \frac{d_k}{|\lambda_2|^k} \cdot \lambda^k ,$$

where d_k are the coefficients of the Chebyshev polynomial $T_p(x) = \sum_{k=o}^{p} d_k x^k$. In this case the quantity (2) equals

$$\frac{1}{|P_p(\lambda_1)|} = \frac{1}{\left|T_p\left(\frac{1}{\gamma}\right)\right|} = \frac{2}{\left[\frac{1}{\gamma} + \sqrt{\frac{1}{\gamma^2} - 1}\right]^p + \left[\frac{1}{\gamma} - \sqrt{\frac{1}{\gamma^2} - 1}\right]^p}$$

$$< \frac{2}{\left[\frac{1}{\gamma} + \sqrt{\frac{1}{\gamma^2} - 1}\ \right]^p} = \frac{2\gamma^p}{\left(1 + \sqrt{1 - \gamma^2}\right)^p} .$$

If, therefore, we choose $P_p(\lambda)$ in the manner indicated, and put

$$\theta_{p+1} = \sum_{k=o}^{p} c_k f_{k+1} = \sum_{i=1}^{n} \alpha_i \lambda_i P_p(\lambda_i) ,$$

the approximate equality $\lambda_1 \approx \frac{\theta_{p+1}}{\theta_p}$ will hold, with an error estimable by the quantity

$$N_p = \frac{2M\left(\frac{\gamma}{1 + \sqrt{1 - \gamma^2}}\right)^p (1 + \gamma)}{1 - 2M\left(\frac{\gamma}{1 + \sqrt{1 - \gamma^2}}\right)^p} .$$

For p sufficiently large, $\frac{N_p}{M_p} \approx 2\left(\frac{1}{1 + \sqrt{1 - \gamma^2}}\right)^p$, and the ratio $\frac{\theta_{p+1}}{\theta_p}$ is significantly closer to λ_1 than the ration $\frac{f_{p+1}}{f_p}$. If $\gamma > 0.5437\cdots$, then $\frac{\gamma}{1 + \sqrt{1 - \gamma^2}} < \gamma^2$, and the method proposed by us thus gives, generally speaking, better results than the method of Aitken.

The application of the method indicated requires an approximate determination of λ_1, then of λ_2, making it possible after this to refine the value of λ_1.

The proper vector X_1 corresponding to the first proper number λ_1 may be refined by the same method. As the approximation to it one usually adopts the vector Y_p (considering the vectors

$Y_o, Y_1, \cdots, Y_p$ to have been found). A better approximation would be the vector $P_p(A)Y_o = \sum_{k=o}^{p} c_k Y_k$. The approximation will in this case have the order $\left(\frac{\gamma}{1 + \sqrt{1 - \gamma^2}}\right)^p$.

2. <u>The solution of a system of linear algebraic equations.</u>

Holding to the assumption of the linearity of the elementary divisors and the reality of the proper numbers of the matrix A, let us consider the system of equations

(3) $$X - AX = Y \quad ,$$

where Y is the given and X the sought vector. The method of iteration will be applicable to this equation if $|\lambda_i| < 1$ $(i = 1,2,\cdots,n)$. The essence of this method consists in the replacement of the exact solution

$$X = (I - A)^{-1}Y = \sum_{k=o}^{\infty} A^k Y = \sum_{k=o}^{\infty} Y_k$$

by a segment of the series standing on the right. Thus one adopts

$$X \approx \sum_{k=o}^{p} Y_k = X^{(p)} \quad .$$

Let us set ourselves the task of finding such a linear combination of the vectors $Y_o, Y_1, \cdots, Y_p$: $Z_p = \sum_{k=o}^{p} c_k Y_k$ as will best approximate the vector X.

Introducing the linearly independent proper vectors of the matrix A: $X_1, \cdots, X_n$ corresponding to the proper numbers $\lambda_1, \cdots, \lambda_n$, let us resolve the vector Y in terms of these vectors:

$$Y = \sum_{i=1}^{n} \alpha_i X_i \quad .$$

Then

$$Y_k = \sum_{i=1}^{n} \alpha_i \lambda_i^k X_i \ , \qquad X = \sum_{i=1}^{k} \alpha_i \frac{1}{1-\lambda_i} X_i \quad .$$

Therefore

$$X - Z_p = \sum_{i=1}^{n} \frac{\alpha_i}{1-\lambda_i} X_i - \sum_{i=1}^{n} \sum_{k=o}^{p} \alpha_i c_k \lambda_i^k X_i = \sum_{i=1}^{n} \alpha_i \left[\frac{1}{1-\lambda_i} - S_p(\lambda_i) \right] X_i \ ,$$

where

$$S_p(\lambda) = \sum_{k=o}^{p} c_k \lambda^k \quad .$$

If we know nothing more about the distribution of the proper numbers than the modulus of the dominant one, λ_1, it will be best to choose as $S_p(\lambda)$ the polynomial of best approximation to the function $\frac{1}{1-\lambda}$ in the interval $[-|\lambda_1|, |\lambda_1|]$.

It was P. L. Chebyshev who found the polynomial of best approximation[1] for the functions $\frac{1}{a-x}$ $(a > 1)$ in the interval $[-1, 1]$. It permits us to write the expression for $S_p(\lambda)$. Let

$$Q_p(\xi) = \frac{2\alpha^{p+1}}{(1-\alpha^2)^2} \cdot \frac{1}{\xi - a} \left[T_{p+1}(\xi) - 2\alpha T_p(\xi) + \alpha^2 T_{p-1}(\xi) \right] - \frac{1}{\xi - a} = \sum_{k=o}^{p} d_k \xi^k \ ,$$

where $a = \frac{1}{|\lambda_1|}$, $\alpha = a - \sqrt{a^2 - 1}$, $T_p(\xi) = \cos p(\text{arc} \cos \xi)$ is the p-th Chebyshev polynomial.

Then

$$S_p(\lambda) = - \sum_{k=o}^{p} \frac{d_k}{|\lambda_1|^{k+1}} \lambda^k \quad .$$

Here for $\lambda \in [-|\lambda_1|, |\lambda_1|]$

- - - - - - - - - - - - - -

[1] N. I. Akhiezer, Lectures on the theory of approximation, Gostekhizdat (1947), p. 69.

$$\left|\frac{1}{1-\lambda} - S(\lambda)\right| \leq \frac{2a\alpha^{p+1}}{(1-\alpha^2)^2} = \frac{\varkappa^p}{2(1-\varkappa^2)(1+\sqrt{1-\varkappa^2})^{p-1}} ,$$

where we have put $\varkappa = |\lambda_1| = \frac{1}{a}$.

Therefore

$$\|X - Z_p\| \leq N \cdot \frac{\varkappa}{2(1-\varkappa^2)} \left(\frac{\varkappa}{1+\sqrt{1-\varkappa^2}}\right)^{p-1} ,$$

where by N is denoted the sum $\sum_{i=1}^{n} |\alpha_i|$, while for the difference $X - X^{(p)}$ one obtains the estimate

$$\|X - X^{(p)}\| \leq N \frac{1}{1-\varkappa} \varkappa^{p+1} .$$

We remark that the method indicated may also be applied in case the Seidel form of the method of iterations is applied to the solution of system (3). Indeed, the Seidel method of iterations is equivalent to the ordinary method of iterations applied to another system

$$X - A_1 X = Y_1 .$$

Lastly we point out that the method here proposed may be applied even when the ordinary method of iterations diverges. It is sufficient to know one or two intervals containing all the proper numbers of the matrix A and not containing the point 1, and to be able to construct the polynomials of best approximation for the function $\frac{1}{1-\lambda}$ in these intervals.

III

SOME ESTIMATES FOR THE METHOD OF STEEPEST DESCENT

by

M. SH. Birman[1]

The present note is devoted to the improvement of certain estimates obtained by L. V. Kantorovich [2] for the generalized method of steepest descent proposed by him.

Let A be a symmetric, bounded, positive definite operator in real Hilbert space; let e_λ be its spectral function; and let the numbers M and m be its upper and lower bounds.

Let us consider the equation

$$(1) \qquad Ax - \phi = 0 \quad .$$

The problem of solving this equation is equivalent to the problem of finding the minimum of the functional

$$(2) \qquad H(x) = (Ax, x) - 2(x, \phi) \quad .$$

The generalized method of steepest descent[2] consists in the construction of a sequence $\{x_n\}$ in which the element x_o is arbitrary, and the subsequent elements are defined by the formula

$$(3) \qquad x_n = x_{n-1} + \sum_{k=o}^{p-1} \alpha_k^{(n)} A^k z_n \quad .$$

Here $z_n = Ax_{n-1} - \phi$, and the numbers $\alpha_k^{(n)}$ are determined from the

- - - - - - - - - - - - - - -

[1] Uspekhi Matematicheskikh Nauk, vol. 5, no. 3 (1950), pp. 152-155.

[2] See L. V. Kantorovich [1], [2].

condition for a minimum of the quantity $H(x_n)$. This condition leads to a system of equations linear in the $\alpha_k^{(n)}$:

$$(A^k z_n, z_n) + (A^k z_n, Az_n)\alpha_0^{(n)} + \cdots + (A^k z_n, A^p z_n)\alpha_{p-1}^{(n)} = 0$$
$$(k = 0, 1, \cdots, p-1) \quad .$$

Let us denote the solution of equation (1) by x^* and put

$$\eta_n = x_n - x^* \; .$$

L. V. Kantorovich has shown that the sequence $\{x_n\}$ is minimizing for the functional (2) and converges strongly to the solution of equation (1); moreover

$$(4) \qquad (A\eta_n, \eta_n) = H(x_n) - H(x^*) \leqq [H(x_0) - H(x^*)] \left(\frac{M-m}{M+m}\right)^{2np} ,$$

and accordingly

$$(5) \qquad \|\eta_n\| \leqq \frac{1}{\sqrt{m}} \left(\frac{M-m}{M+m}\right)^{np} [H(x_0) - H(x^*)]^{\frac{1}{2}} \quad .$$

It was L. V. Kantorovich's surmise [2] that for $p > 1$ these estimates could be improved.

Indeed, as it has turned out, estimates (4) and (5) can be replaced by the following:

$$(4a) \qquad H(x_n) - H(x^*) \leqq L_p^{2n}[H(x_0) - H(x^*)]$$

and

$$(5a) \qquad \|\eta_n\| \leqq \frac{L_p^n}{\sqrt{m}} \cdot [H(x_0) - H(x^*)]^{\frac{1}{2}} \quad .$$

Here L_p is a quantity expressible simply in terms of the deviation from zero of the p-th Chebyshev polynomial, viz.:

$$L_p = \frac{2}{\left[\frac{M+m}{M-m} - \sqrt{\left(\frac{M+m}{M-m}\right)^2 - 1}\right]^p + \left[\frac{M+m}{M-m} + \sqrt{\left(\frac{M+m}{M-m}\right)^2 - 1}\right]^p} \quad .$$

In particular,

$$L_1 = \frac{M-m}{M+m} \; ; \; L_2 = \frac{(M-m)^2}{(M+m)^2+4mM} \; ; \; L_3 = \frac{(M-m)^3}{(M+m)[(M+m)^2+12mM]} \; ; \; \cdots \text{, etc.}$$

The idea of the proof consists in constructing a majorant sequence of definite form as close as possible to the sequence $\{x_n\}$.[1]

Let us replace equation (1) by the equivalent equation

(6) $$x = x + \sum_{k=0}^{p-1} \varepsilon_k A^k (Ax - \emptyset) \quad .$$

We shall try to choose the numbers ε_k so that the norm of the operator

$$B_p = I + A \sum_{k=0}^{p-1} \varepsilon_k A^k$$

will be less than unity for any symmetric operator A with bounds[2] M and m.

Let λ be a spectral point of the operator A, and μ a spectral point of the operator B_p. Then

(7) $$\mu = \mu(\lambda) = 1 + \lambda \sum_{k=0}^{p-1} \varepsilon_k \lambda^k$$

- - - - - - - - - - - - -

[1] I consider it fitting to note that in the proof I have utilized the idea of the application of polynomials of best approximation in the operator variable, which idea has been applied by M.K. Gavurin in hastening the convergence of iterative processes (see the note by M. K. Gavurin appearing in the present issue [the preceding translation]).

[2] I. P. Natanson [3] has proposed such a transformation for the case p = 1, with a view to making the method of successive approximations applicable.

and

(8) $$\|B_p\| \leqq \max_{m \leqq \lambda \leqq M} |\mu| \quad .$$

We shall choose the numbers ε_k so that the function (7) will be a polynomial in λ of degree p deviating least from zero in the interval [m, M] on condition that $\mu(0) = 1$. The ordinary Chebyshev polynomial $T_p(\lambda)$ for the interval [m, M], normalized by the condition $T_p(0) = 1$, is such a polynomial. The maximum of its modulus in [m, M] is L_p. Obviously $1 > L_1 > L_2 > \cdots$, and therefore from (8) it follows that

$$\|B_p\| \leqq L_p < 1 \quad ,$$

and for equation (6) the usual successive approximations, computed by the formula

(9) $$X_n = X_{n-1} + \sum_{k=o}^{p-1} \varepsilon_k A^k (AX_{n-1} - \phi)$$

converge.

Now put

$$\theta_n = X_n - x^* \quad .$$

It follows from (9) that

$$\theta_n = B_p \theta_{n-1} \quad .$$

Hence

$$(A\theta_n, \theta_n) = (AB_p\theta_{n-1}, B_p\theta_{n-1}) = (AB_p^2\theta_{n-1}, \theta_{n-1}) =$$

(10) $$= \int_m^M \lambda\mu^2(\lambda) d(e_\lambda\theta_{n-1}, \theta_{n-1}) \leqq \max_{\lambda \varepsilon [m,M]} \mu^2(\lambda) \int_m^M \lambda d(e_\lambda \theta_{n-1}, \theta_{n-1})$$

$$= L_p^2 (A\theta_{n-1}, \theta_{n-1}) \quad .$$

Inequality (10) will remain true if the quantities η_n and η_{n-1} be substituted for the quantities θ_n and θ_{n-1}, since the quantities $\alpha_k^{(n)}$ are determined at each step from the condition for a minimum

of the expression[1]

$$(A\eta_n, \eta_n) = H(x_n) - H(x^*) \quad .$$

From the inequality

$$(A\eta_n, \eta_n) \leqq L_p^2(A\eta_{n-1}, \eta_{n-1})$$

there follows directly the inequality

$$(4a) \qquad H(x_n) - H(x^*) \leqq L_p^{2n}[H(x_o) - H(x^*)] \quad .$$

Estimate (4a) is exact; to wit: for each p one can find a positive definite symmetric operator A with bounds M and m and an initial element x_o such that for each n we will have

$$(11) \qquad (A\eta_n, \eta_n) = H(x_n) - H(x^*) = L_p^{2n}[H(x_o) - H(x^*)] \quad .$$

For example, with $p = 1$ equation (11) will be satisfied for an operator with two proper numbers $\lambda_o = m$ and $\lambda_1 = M$, if we put

$$x_o = x^* + \gamma_o y_o + \gamma_1 y_1 \quad .$$

Here y_o and y_1 are proper elements of the operator concerned, and the numbers γ_o and γ_1 are connected by the relation

$$\gamma_1^2 \cdot \lambda_1^2 = \gamma_o^2 \cdot \lambda_o^2 \quad .$$

Proceeding from (4a), $\|x_n - x^*\|$ may be easily estimated, to wit

$$(\eta_n, \eta_n) \leqq \frac{1}{m}(A\eta_n, \eta_n) \leqq \frac{L_p^{2n}}{m}(A\eta_o, \eta_o) \quad ,$$

whence (5a) also follows, as well as the inequality

$$\|x_n - x^*\| \leqq \sqrt{\frac{M}{m}} \cdot L_p^n \cdot \|x_o - x^*\| \quad .$$

---

[1]The functionals H(x) and H(x)-H(x^*) attain a minimum simultaneously.

The method of steepest descent is also extended by L. V. Kantorovich to a series of problems in which A is an unbounded operator. It can be shown that in this case as well estimates analogous to (4a) and (5a) hold.

CITED LITERATURE

(all in Russian)

[1] L. V. Kantorovich, "The method of steepest descent," Akad. Nauk SSSR, Doklady, vol. 56 (1947), p. 233.

[2] L. V. Kantorovich, "Functional analysis and applied mathematics," Uspekhi mat. nauk, vol. 3, no. 6 (1948), p. 89. [An English translation is in preparation at the National Bureau of Standards, Los Angeles.]

[3] I. P. Natanson, Leningad, Leningradskiĭ pedagogicheskii institut imeni A. I. Gertsena, Uchënye zapiski, vol. 64 (1948).

IV

A PROCESS OF SUCCESSIVE APPROXIMATIONS FOR FINDING CHARACTERISTIC VALUES AND CHARACTERISTIC VECTORS

by

N. Azbelev and R. Vinograd[1]

(Submitted by Academician I. G. Petrovsky, June 27, 1951)

Let there be given a linear operator A in n-dimensional space (real or complex). The operator is not required to be symmetric.

The process of successive approximations described below leads to any characteristic vector of the operator A, provided a suitable zero-th approximation be chosen. The character of the convergence is specified by Theorem 2.

The process is defined as follows:

An arbitrary vector $x_0 \neq 0$ is adopted as the zero-th approximation. We then put $\lambda_0 = \frac{(Ax_0, x_0)}{\|x_0\|^2}$ and construct the operator $B_0 = A - \lambda_0 E$ and the vector $y_0 = B_0^* B_0 x_0$, afterwards taking as the next approximation $x_1 = x_0 - \eta_0 y_0$, η_0 being a numerical coefficient equalling $\frac{(y_0, x_0)}{\|y_0\|^2} = \frac{\|B_0 x_0\|^2}{\|y_0\|^2}$. We thereby consider λ_0 to be the zero-th approximation to the characteristic value. Carrying the same operations through on x_1 we obtain λ_1 and x_2. In general, supposing λ_{k-1} and x_k to have been already determined, we put

- - - - - - - - - - - - - -

[1]Akademiia Nauk SSSR, Doklady, vol. 83 (1952), pp. 173-174.

$\lambda_k = \dfrac{(Ax_k, x_k)}{\|x_k\|^2}$, and then $B_k = A - \lambda_k E$, after which we determine

$y_k = B_k^* B_k x_k$ and $\eta_k = \dfrac{(y_k, x_k)}{\|y_k\|^2} = \dfrac{\|B_k x_k\|^2}{\|y_k\|^2}$ and lastly $x_{k+1} = x_k - \eta_k y_k$.

The process of the approximations is definite.

Theorem 1 gives a rough estimate of the convergence.

Theorem 1. Depending on the choice of x_o, one of two cases obtains: either $x_k \to 0$, or all the limiting vectors for the sequence $\{x_k\}$ belong to the same characteristic subspace and have the identical norm $d > 0$.

This result is refined by Theorems 2 and 3.

Theorem 2. In the Jordan normal form of the matrix (A) of the operator A let there correspond to the characteristic value λ' only a diagonal box, i.e., the invariant subspace L' belonging to λ' consists of characteristic vectors only. The process then converges with the rapidity of a geometrical progression to a non-zero vector x' of L' (i.e., $\|x_k - x'\| \leq Cq^k$, where $q < 1$) if for x_o an arbitrary vector be chosen that forms with L' an angle less than a certain α_o.

Observation: Examples show that if the condition that the box be diagonal be violated, convergence with a rapidity of only 1/k is observed.

Theorem 3. In the conditions of Theorem 2, as regards L', the "region of attraction" G(L') of the subspace L', i.e., the set of zero-th approximations such as will lead to non-zero vectors from L', is an open set.

Observation 1: Theorem 2 established only that G(L') contains

a certain sufficiently narrow "cone" around L'.

Observation 2: In case operator A is symmetric, the number q figuring in the estimate of the convergence (Theorem 2) is simply expressible in terms of the characteristic values of the operator A. Let the latter be arrayed in ascending order, $\lambda_1 < \lambda_2 < \dots < \lambda_r$ ($r \leqq n$). If we are located in the "region of attraction" of the subspace L_p belonging to λ_p,

$$q = \frac{M_p - m_p}{M_p + m_p} \quad ,$$

where M_p is the larger of the two numbers $(\lambda_1 - \lambda_p)^2$ and $(\lambda_r - \lambda_p)^2$, and m_p is the lesser of the two numbers $(\lambda_{p-1} - \lambda_p)^2$ and $(\lambda_{p+1} - \lambda_p)^2$ (in case p = 1 or p = r we adopt $\lambda_{p-1} = \lambda_p$ or $\lambda_{p+1} = \lambda_p$ respectively).

Observation 3: Let the Jordan normal form of the matrix A be diagonal. The sum of the "regions of attraction" of all the characteristic subspaces is, by Theorem 3, an open set, and consequently the complementary set F -- which is then, on the strength of Theorem 1, a "region of attraction" of zero -- is closed.

The validity of the following hypothesis appears to be very probable: "The 'region of attraction' of zero, at least in case A is reducible to diagonal form, is a nowhere dense set."

Were it proved, this proposition would imply that the case of the convergence of the process to zero is practically impossible. The hypothesis is easily proved for all operators in two-dimensional space (even without assuming A to be reducible to diagonal form) and for some operators in three-dimensional space. We have not succeeded