QUARTERLY OF APPLIED MATHEMATICS

Vol. XIII APRIL, 1955 No. 1

DETERMINATION OF THE ROOT SYSTEMS OF ALGEBRAIC EQUATIONS BY AFFINITY TRANSFORMS*

BY

C. A. TRAENKLE

Aeronautical Research Laboratory, Wright Field

1. Review of methods and results. The problem to be analysed here, one with a wide range of applications, is to evaluate the roots of an nth order equation

$$a_n z^n + a_{n-1} z^{n-1} + \cdots a_2 z^2 + a_1 z + a_0 = 0; \tag{1}$$

the roots z may be real or complex numbers.

There are various ways of solving the problem:

a. In one group of methods, estimated approximations of z are substituted in Eq. (1) and varied by trial and error until the functional value of this equation becomes zero. If z is complex, the computational labor increases considerably. To alleviate it, several electronic mechanical computers of the analogue type have been developed, e.g. by H. C. Hart and J. Travis [1]. A modification of this group by W. R. Evans [2] applies the special concepts of the theory of servomechanisms.

b. The well known root squaring method of Graeffe transforms the coefficients a_i of Eq. (1) gradually into a sequence, identical with the higher squared root values. However the serious disadvantage of this method is that it yields only the moduli of the complex roots, whereas the phase angles have to be determined separately by additional methods given by F. W. J. Olver in [3].

c. The process of factorization splits Eq. (1) into separate factors giving immediately the root values. For linear factors, i.e. real roots, the process is identical with the well-known *Regula Falsi*, Newton's rule and Horner's method. For quadratic factors, which include conjugate complex roots, the procedure is due to L. Bairstow [4]: starting with an estimated approximation of the root factor, the coefficients a_i of Eq. (1) are transformed by long division into the coefficient set b_i. The remainder terms of b_i, together with an additional coefficient set c_i derived from the b_i, are used to calculate corrections for the root parameters. The b_i- and c_i-sets are computed by reduction formulae suitable for routine work. The method represents an extension to Newton's rule for the complex. It was later systematized by W. E. Milne [5]. The procedure by S. N. Lin [6] corresponds analogously to a simplified correction formula; however it is convergent only for restricted cases.

The above factorization process is extended here by the following results.

*Received Jan. 15, 1954. The author acknowledges the encouragement given by Col. L. B. Williams, Dr. M. G. Scherberg, Dr. J. E. Clemens and B. B. Johnstone of the Aeronautical Research Laboratory, Wright Field.

In simplifying the procedure, the transform matrix and the corrections for the root values are derived from the b_i-sets alone. As the convergence speed depends largely on the first good approximations of the root parameters, a system of rules of calculated estimates is established. The resulting procedure is simple and converges rapidly.

The variations of the a_i-coefficients are linearly dependent on the increments of the root parameters and vice versa. The matrices of those nth order systems are computed by direct procedures, giving important information about the properties of the root systems with respect to the intended application. A special case of these linear systems is the error relation between the variations of the a_i-coefficients and the variations of the root parameters.

Multiple roots represent singularities as the linear relations of the transform matrix cease to be valid. It is shown how the error range of the multiple roots can nevertheless be computed directly, making the case practically identical with that of near multiple roots. Numerical examples illustrate the details and the specific relations in this transition range.

2. Reduction formulae for factorization. A pair of conjugate complex roots p, p', put into factor form as $(z + p) = 0$, $(z + p') = 0$, are composed into the resultant root factor by multiplying

$$(z + p)(z + p') = z^2 + p_1 z + p_0 , \tag{2}$$

where $p_1 \equiv p + p'$, $p_0 \equiv pp'$ are real numbers again. Now factor (2) is divided into Eq. (1) by long division

$$\begin{aligned}(a_n z^n + a_{n-1} z^{n-1} + \cdots a_2 z^2 + a_1 z + a_0) &\div (z^2 + p_1 z + p_0) \\ &= b_n z^{n-2} + b_{n-1} z^{n-3} + \cdots b_2 + (b_1 z + b_0)/(z^2 + p_1 z + p_0). \end{aligned} \tag{3}$$

The remainder terms b_1 , b_0 would be zero if factor (2) were already correct; otherwise they can be used to compute corrections for the root parameters p_1 , p_0 to remove the remainder b_1 , b_0 finally. For numerical computation the a_i-coefficients shall be divided by a_n before being processed; therefore $a_n = b_n = 1$. The reduction formulae for the b_i-coefficient set are

$$b_i = -p_0 b_{i+2} - p_1 b_{i+1} + a_i , \tag{4}$$

with the remainder coefficients

$$b_1 = p_0 b_3 - p_1 b_2 + a_1 , \tag{5}$$

$$b_0 = -p_0 b_2 \qquad + a_0 . \tag{5'}$$

The reduction formulae b_i are identical with those of W. E. Milne [5], but the remainder term b_0 in (5′) is different, as the term $-p_1 b_1$ cannot enter it, b_2 being the last term of the quotient.

At first as an approximation for the general case, the correction formulae for the root parameters are deduced for the special case of the lowest of widely separated roots. According to the idea of the Graeffe method, the terms in Eq. (1) higher than z^2 can be neglected here:

$$a_2 z^2 + a_1 z + a_0 = 0.$$

When compared with the root factor $z^2 + p_1 z + p_0$, this gives a first approximation for the root parameters

$$p_0 = a_0/a_2 , \qquad p_1 = a_1/a_2 .$$

The division form of Eq. (3) with the approximate parameters p_0, p_1 may be rearranged by multiplying both sides by the quadratic factor

$$a_n z^n + \cdots a_0 = (b_n z^{n-2} + \cdots b_2)(z^2 + p_1 z + p_0) + b_1 z + b_0 = 0.$$

Here again the terms higher than b_2 are neglected according to the above idea. After simplifying and rearranging one obtains the new remainder

$$z^2 + (p_1 + b_1/b_2)z + (p_0 + b_0/b_2) = 0,$$

which, when compared with the starting root factor, gives the corrections

$$\Delta p_0 = b_0/b_2 , \qquad \Delta p_1 = b_1/b_2 ; \tag{6}$$

they show a linear interdependence on the b_1, b_0. The general form would be

$$\left.\begin{aligned} \Delta p_0 &= k_{01} b_1 + k_{00} b_0 \\ \Delta p_1 &= k_{11} b_1 + k_{10} b_0 \end{aligned}\right\}, \tag{7}$$

representing a linear affine relationship with the transform matrix

$$K \equiv \begin{bmatrix} k_{01} & k_{00} \\ k_{11} & k_{10} \end{bmatrix}. \tag{8}$$

The formulae are so arranged to conform with the computation forms of Sec. 6 below, embodying the rules of matrix multiplication—column $b_1 b_0$ multiplied by row $k_{01} k_{00}$; therefore the principal and secondary diagonal of K had to be interchanged.

The special correction formulae (6) with their matrix

$$K = \begin{bmatrix} 0 & 1/b_2 \\ 1/b_2 & 0 \end{bmatrix} \tag{6'}$$

are identical with those of S. N. Lin [6], but their application is confined to the restricted range of widely separated roots in consequence of the supposition made above.

Having computed the parameters p_0, p_1 of the quadratic factor (2), the roots can be resolved by the well-known formula $z = -p_1/2 \pm (p_1^2/4 - p_0)^{1/2}$; they may turn out as conjugate complex or real. In this case the two linear real factors $(z + p)$ could be alternately split off from the original equation quite as well. If the degree of the given equation is an odd number, this is even necessary for the one linear factor. Then the above general formulae have to be modified and become the well-known Horner's scheme for real roots

Division form

$$z^n + \cdots a_0 = (b_n z^{n-1} + \cdots b_1)(z + p) + b_0 = 0,$$

Reduction formula

$$b_i = -p b_{i+1} + a_i , \tag{4'}$$

Remainder

$$b_0 = -p b_1 + a_0 ,$$

p-correction

$$\Delta p = k b_0 , \tag{7'}$$

Transform coefficient

$$k \equiv 1/b_1 .$$

Equation (7′) is plainly identical with Newton's rule for real roots, whereas Eq. (7) is an analogous generalized form.

3. Affinity transform of remainder into root correction. The remainder coefficients b_1 , b_0 and the corrections Δp_0 , Δp_1 are interpreted as the components of a vector in

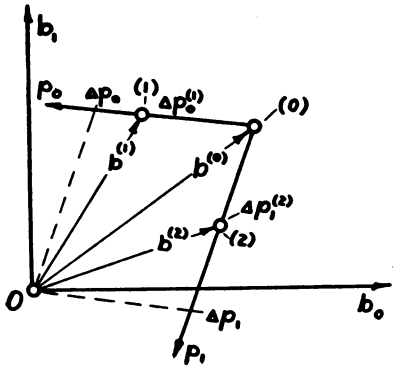

FIG. 1. Graphical root solution

the b- and p-planes respectively (see Fig. 1). The problem is now to establish the transformation from one plane into the other.

This can be done illustratively by graphical means. The starting values of b_1 , b_0 and p_1 , p_0 are represented by the point $^{(0)}$ or vector $b^{(0)}$ of Fig. 1. Now two increment vectors $\Delta p^{(1)}$ and $\Delta p^{(2)}$ are arbitrarily chosen as probing steps, and their b-vectors are computed according to Eqs. (4), (5) giving the points (1) and (2). The p-increments may be chosen alternately along the p-axes

$$\left.\begin{matrix} \Delta p_0^{(1)} \\ \Delta p_1^{(1)} = 0 \end{matrix}\right\}, \quad \left.\begin{matrix} \Delta p_0^{(2)} = 0 \\ \Delta p_1^{(2)} \end{matrix}\right\}.$$

Then (0) (1) and (0) (2) are the images of the p_0- and p_1-axes respectively, whereby the distance (0) (1) divided by $\Delta p_0^{(1)}$ is the scale value of p_0 ; that of p_1 is obtained correspondingly. The vector $(0)0 \equiv -b^{(0)}$ is hereafter decomposed along the p_0- and p_1-axes, and these components ratioed by the scale factors are immediately the wanted corrections Δp_0 , Δp_1 . Of course these computations are only applicable within the linear range.

The analytical form of this approach has some advantages with regard to more general application. For this purpose the increment vectors are referred to the starting point (0)

$$\left.\begin{matrix} \Delta p^{(1)} = p^{(1)} - p^{(0)} \\ \Delta p^{(2)} = p^{(2)} - p^{(0)} \end{matrix}\right\} \quad \left.\begin{matrix} \Delta b^{(1)} = b^{(1)} - b^{(0)} \\ \Delta b^{(2)} = b^{(2)} - b^{(0)} \end{matrix}\right\}. \tag{9}$$

These increment vectors have to comply with Eq. (7), whereby the Δb-vectors have to be supplemented by a negative sign, as the latter and the zero vector $b^{(0)}$ are of opposite sign by definition:

Point (1)

$$\left.\begin{aligned} \Delta p_0^{(1)} &= -(k_{01}\Delta b_1^{(1)} + k_{00}\Delta b_0^{(1)}) \\ \Delta p_0^{(2)} &= -(k_{01}\Delta b_1^{(2)} + k_{00}\Delta b_0^{(2)}) \end{aligned}\right\}, \tag{10}$$

Point (2)

$$\left.\begin{aligned} \Delta p_1^{(1)} &= -(k_{11}\Delta b_1^{(1)} + k_{10}\Delta b_0^{(1)}) \\ \Delta p_1^{(2)} &= -(k_{11}\Delta b_1^{(2)} + k_{10}\Delta b_0^{(2)}) \end{aligned}\right\}. \tag{11}$$

These two pairs of linear equations can be used to compute the two pairs of unknown coefficients k_{01}, k_{00} and k_{11}, k_{10} of the matrix K. According to Cramer's rule

$$\begin{aligned} k_{01} &= -D^{-1}\begin{vmatrix} \Delta p_0^{(1)} & \Delta b_0^{(1)} \\ \Delta p_0^{(2)} & \Delta b_0^{(2)} \end{vmatrix} = -D^{-1}\begin{vmatrix} \Delta p_0^{(1)} & \Delta p_0^{(2)} \\ \Delta b_0^{(1)} & \Delta b_0^{(2)} \end{vmatrix}, \\ k_{00} &= -D^{-1}\begin{vmatrix} \Delta b_1^{(1)} & \Delta p_0^{(1)} \\ \Delta b_1^{(2)} & \Delta p_0^{(2)} \end{vmatrix} = +D^{-1}\begin{vmatrix} \Delta p_0^{(1)} & \Delta p_0^{(2)} \\ \Delta b_1^{(1)} & \Delta b_1^{(2)} \end{vmatrix}, \end{aligned} \tag{10'}$$

$$\begin{aligned} k_{11} &= -D^{-1}\begin{vmatrix} \Delta p_1^{(1)} & \Delta b_0^{(1)} \\ \Delta p_1^{(2)} & \Delta b_0^{(2)} \end{vmatrix} = -D^{-1}\begin{vmatrix} \Delta p_1^{(1)} & \Delta p_1^{(2)} \\ \Delta b_0^{(1)} & \Delta b_0^{(2)} \end{vmatrix}, \\ k_{10} &= -D^{-1}\begin{vmatrix} \Delta b_1^{(1)} & \Delta p_1^{(1)} \\ \Delta b_1^{(2)} & \Delta p_1^{(2)} \end{vmatrix} = +D^{-1}\begin{vmatrix} \Delta p_1^{(1)} & \Delta p_1^{(2)} \\ \Delta b_1^{(1)} & \Delta b_1^{(2)} \end{vmatrix}, \end{aligned} \tag{11'}$$

where

$$D = \begin{vmatrix} \Delta b_1^{(1)} & \Delta b_0^{(1)} \\ \Delta b_1^{(2)} & \Delta b_0^{(2)} \end{vmatrix} = \begin{vmatrix} \Delta b_1^{(1)} & \Delta b_1^{(2)} \\ \Delta b_0^{(1)} & \Delta b_0^{(2)} \end{vmatrix}. \tag{12}$$

The above determinants have been transformed by mirroring at the diagonal and by interchanging rows. These computations are easily and rather mechanically performed in the special computation form of Sec. 6 below.

The above transform matrix

$$K = \begin{bmatrix} k_{01} & k_{00} \\ k_{11} & k_{10} \end{bmatrix}$$

for the Δp-corrections (7) is computed by three positions of the remainder b_1, b_0. Therefore this approach can be regarded as a generalized adaptation of the *Regula Falsi* to complex roots.

In order to compute the remainder vector b_1, b_0 according to Eqs. (4), (5), the co-

efficients a_i and b_i are arranged in columns, see Fig. 2. The actual root parameters p_0, p_1 are written on a paper strip with reversed sign, multiplied and summed up with the other coefficients by a desk computer with accumulative multiplication according to Milne [5], as shown in Fig. 2. This procedure is used in Table 1 and works rather mechanically. The dotted line below $i = 2$ serves as a reminder that in Eq. (5′) for b_0 the routine term $-p_1 b_1$ must *not* be included.

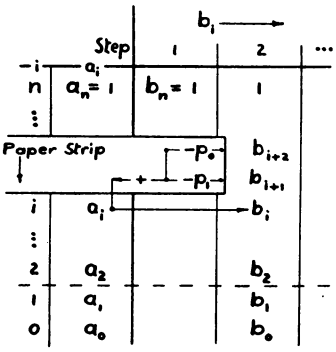

FIG. 2. b_i-Paper strip plot

4. General computation procedure with p-traces. The affinity transform of Eqs. (7), (8) should only be applied if the functions b_1, b_0 of the remainder vector are sufficiently linear. If this is not yet the case, one starts with some approximate values p_0, p_1, images the corresponding point into the b_1, b_0-plane (point 1 in Fig. 3) and subsequently additional points 2, 3 for varying p_0 but p_1 = constant. In the b_1, b_0-plane they constitute a curve (the p_0-trace), possibly showing both considerable curvature and non-uniformity in the p_0-scale. In the region of closest proximity to the origin 0 (point 2), one starts

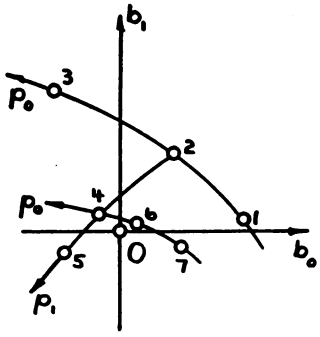

FIG. 3. p-Traces

the next trace p_1 (p_0 = constant), imaging additional points 4, 5 and selecting again the region of closest proximity to 0. With these successive steps one can always approximate any root value as accurately as necessary.

During approximation one will reach a region with sufficient linearity (e.g. points 4, 5, 6). When this happens one should switch over to the affinity transform of Sec. 3, as this assures the most rapid convergence.

As shown in Fig. 3, the images of the p-coordinate axes (p-traces) may be rotated with reference to the b-axes by any amount and they may have any obliqueness. But the sign of both systems remains the same (right hand system), following at first from Eq. (6) for the limit of widely separated roots, showing that both the p- and b-axes are hereby even coincident and, of course, orthogonal. The other limit is that of multiple roots, where the p-traces coincide with each other, as shown in Sec. 10 below. As all the actual cases lie between these two limits the above statement of the sign invariance of systems is proven. It can be used as a shortcut in selecting the subsequent steps always on the correct side of the p-traces.

5. Shortcut procedure with calculated estimates for the p-steps. For a quick convergence one has to start with a good approximation of the root parameters p_0, p_1, one which already falls into the linearity range of K. A well-known approximation for the lowest and highest root factor exploits the idea of Graeffe's method and neglects alternately the higher and the lower terms of z^i; by taking only the first or last three terms of Eq. (1), one obtains

$$G\text{-Max.} \quad \left.\begin{aligned} p_0 &= a_{n-2} \\ p_1 &= a_{n-1} \end{aligned}\right\} \text{ as } a_n = 1, \qquad G\text{-Min.} \quad \left.\begin{aligned} p_0 &= a_0/a_1 \\ p_1 &= a_1/a_2 \end{aligned}\right\}. \tag{13}$$

These formulae are fairly accurate if the root values are sufficiently separated. If however they reverse their turn or contradict each other, the roots must lie close together. Another approximation is furnished by the mean root $p_{0m}p_{1m}$. By multiplying the $h = n/2$ quadratic factors one obtains

$$(z^2 + p_{11}z + p_{01})(z^2 + p_{12}z + p_{02}) \cdots (z^2 + p_{1h}z + p_{0h}) = 0,$$

or

$$z^n + \sum p_{1h}z^{n-1} + \cdots p_{01}p_{02} \cdots p_{0h} = 0.$$

Comparing with Eq. (1)

$$\sum p_{1h} = a_{n-1}, \qquad \prod p_{0h} = a_0,$$

and by definition

$$\left.\begin{aligned} p_{0m} &\equiv (\prod p_{0h})^{1/h} = a_0^{1/h}, \\ p_{1m} &\equiv 1/h \sum p_{1h} = a_{n-1}/h \end{aligned}\right\}. \tag{14}$$

The normal procedure is always to eliminate the lowest factor in the successive elimination process. If the G-values of Eq. (13) get entangled, the minimum values can be derived accurately enough from the mean value and the maximum value already known by G-Max. for $h = 1$. These shall be termed as Estimated Min. (E-Min.) to distinguish them from the Graeffe Min. (G-Min.).

E-Min.

$$\left.\begin{aligned} p_{0\text{Min}} &\approx p_{0m}^2/p_{0\text{Max}} \approx p_{0m} - (p_{0\text{Max}} - p_{0m}), \\ p_{1\text{Min}} &\approx p_{1m} - (p_{1\text{Max}} - p_{1m}). \end{aligned}\right\}. \tag{15}$$

In numerical computation, the values (13), (14), (15) are compiled in tables (see Estimation Section in Table 1). The E-Min.-values are then taken as the starting steps in the computation section of the tables; these values are already rather good approximations as proven by experience.

The next step 2 is taken on a p_0-trace with a roughly estimated Δp_0 ; the b-remainder vector of step 1 is plotted in a b_1 , b_0-diagram together with a fictitious step o with $p_0 = 0$, $p_1 = 0$, which gives $b_1 = a_1$, $b_0 = a_0$; the origin is projected on (o 1) as 0′, and by linear interpolation one takes roughly $\Delta p_0 \approx p_0(1\ 0')/(o\ 1)$, including the correct sign (see Fig. 4). After having the remainder vector of step 2 subsequently transferred to the diagram, one can already roughly plot a first pair of p_0p_1-traces according to Sec. 4, from which the corrections for step 3 can be derived (Fig. 4). After having thus computed

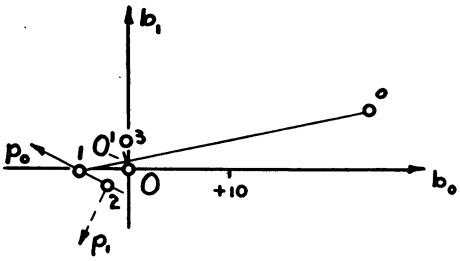

FIG. 4. Remainder-root-diagram

3 points as preliminary steps near the origin and obviously within the linearity range, the transform matrix K can now be determined (see Table 2).

6. Numerical example, convergence number. The above outlined method will be further clarified by means of a numerical example. For convenience of comparison the algebraic equation to be solved is taken from Lin's paper [6] (Hitchcock's example)

$$z^8 - 3.012z^7 + 3.225z^6 + 1.021z^5 + 6.986z^4 - 21.887z^3 + 8.110z^2 + 5.901z + 23.889 = 0.$$

It is assumed that the a_i-coefficients are correct to 1/2 unit of the last given digit.

Table 1 gives the solution of the lowest root couple ($h = 1$). It consists of the esti-

Estim.	G-Max.	G-Min.	Mean	E-Min.
p_0	+3.2	+3	+2.2	+1
p_1	-3	+1	-1	+1

Computation	Step	Preliminary Steps 1	2	3	K - Steps 4	5
	p_0	+1	+0.9	+0.9	+0.861	+0.8605
	p_1	+1	+1.0	+0.9	+0.945	+0.9430
i	a_i	b_i				
8	+1	+1	+1	+1	+1	+1
7	-3.012	-4.012	-4.012	-3.912	-3.957	-3.955
6	+3.225	+6.237	+6.337	+5.846	+6.103	+6.094
5	+1.021	-1.204	-1.705	-0.720	-1.339	-1.322
4	+6.986	+1.953	+2.988	+2.373	+2.997	+2.989
3	-21.887	-22.636	-23.341	-23.375	-23.566	-23.568
2	+8.110	+28.793	+28.762	+27.012	+27.799	+27.763
1	+5.901	-0.256	-1.854	+2.628	-0.079	+0.001
0	+23.889	-4.904	-1.997	-0.422	-0.046	-0.001
	β	(0.6)	(0.9)	0.03	0.02	-

TABLE 1 *Root h* = 1

mation- and computation section. After the preliminary steps have been set up, the K-matrix is computed in Table 2 according to Eqs. (10′), (11′), (12); the sequence of operations is run mechanically along its symbolical sequence-circuit. By some K-steps more the remainder vector b_1 , b_0 is finally put to zero or below the error limit. The process converges rapidly.

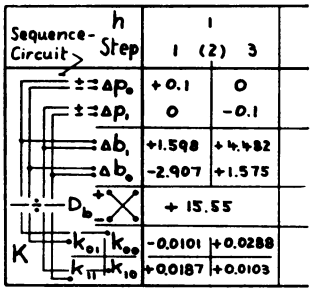

TABLE 2 *Transform matrix* K_h

After the lowest root is thus eliminated, the b_i-coefficients of the last step are taken as the a_i-coefficients of a new equation whose degree is reduced by two. These new abridged coefficients are again processed according to the outlined method, eliminating the next lowest root. This process is continued, until the last root factor has been extracted.

Obviously the abridged a_i-coefficients already contain some errors from the former root factors, which reflect again in additional errors of the new root parameters. In fact this error propagation is rather small and may usually be disregarded. But for a

i	a_i (error a_a ±.0005)	h: b_{ih} 1	error a_b	2		3		4	
8	+1	+1		+1		+1		+1	
7	-3.012	-3.955		-5.098		+0.007		+0.010	
6	+3.225	+6.094		+11.623		+0.591		-1.426	
5	+1.021	-1.322		-11.831		+2.786		-3.335	
4	+6.986	+2.989		+5.697		+13.827		+3.583	
3	-21.887	-23.568		-7.341		+12.454		+4.553	
2	+8.110	+27.763		+10.694		+8.993		+5.096	
1	+5.901	+0.001	±.001	-0.004	±.008	-0.018	±.019	-0.012	±.058
0	+23.889	-0.001	±.001	+0.002	±.013	+0.013	±.042	+0.034	±.164
	p_{0h}	+0.8605	error Δp ±.00003	+2.2337	±.0001	+2.655	±.0006	+4.681	±.0009
	p_{1h}	+0.9430	±.00002	+2.0855	±.0001	-3.019	±.0004	-3.022	±.0002
K_h	k_{01} / k_{00}	-.0101	+.0288	-.0017	-.0102	+.0441	+.0327	-.0177	-.0010
	k_{11} / k_{10}	+.0187	+.0103	-.0114	+.0004	-.0176	-.0168	+.0104	+.0040

TABLE 3 *Unabridged* b_i*-coefficients, p-root parameters and K-matrices*

closer analysis this influence can be eliminated by going back to the original a_i-coefficients for each root factor in a second round, computing the unabridged b_i-sets, K-matrices and root corrections Δp as above. These results are compiled in Table 3 and moreover are later required for the system analysis of Sec. 8 below.

The convergence of the computation can be checked by the "convergence number" defined for the qth step as

$$\beta_q \equiv | b_{1,0}^{(q+1)} | / | b_{1,0}^{(q)} |, \tag{16}$$

i.e. the ratio of the absolute amounts of two subsequent remainder vectors. For the first K-step the denominator is the mean of the $| b_{1,0} |$ of all three preliminary steps as they all contribute equally to it. The convergence number is characteristic for the deviation from linearity, the error ΔK of K and the error $\Delta(\Delta p)$ of the correction Δp, namely

$$\beta \approx \Delta K / K \approx \Delta(\Delta p)/\Delta p. \tag{16'}$$

For the preliminary steps the following order of magnitude should normally be attained

$$\beta_{pr} \lesssim 0.1 \cdots 1,$$

and for the K-steps

$$\beta_K \lesssim 0.01 \cdots 0.1.$$

The convergence of the K-steps is very rapid; for example $\beta_k = 0.01$ indicates a gain of two digits per step. It is remarkable that these limits are usually reached by means of the first K-matrix, as for example in Table 1. If this should not happen, however, one has simply to compute the subsequent K-matrix.

It is instructive to compare the exact results of Table 3 with those of Lin [6], according to the simplified form (6′); only the lowest and most separated root couple $h = 1$ has a sufficient overall accordance between the simplified K and the complete form to give convergence. For the other roots the secondary diagonals of K take over in importance, or even the signs of the principal diagonals reverse, thus changing bluntly from convergence to divergence. In order to overcome these difficulties, Lin has to increase the separation of the different roots by supplementary methods, which makes the procedure tedious and the results inaccurate. This lack of convergence is also investigated in a paper by F. B. Hildebrand [7].

7. Synthetic multiplication. As a final check all the extracted root factors are multiplied by each other; the resulting coefficients should be identical with the original a_i. The computation is compiled in Table 4 and is executed by means of a paper strip plot according to Fig. 5. As a rough estimate the mean relative deviation of the a_i may be

h	4	3	2	1		
p_0	+4.681	+2.655	+2.2337	+0.8605	original	$\Delta a_i/a_i$
p_1	−3.022	−3.019	+2.0855	+0.9430	Coeff. a_i	10^{-4}
i	b_i			a_i		
8				+1	+1	–
7				−3.012	−3.012	0
6			+1	+3.225	+3.225	0
5			−3.955	+1.019	+1.021	20
4		+1	+6.094	+6.983	+6.986	4
3		−6.041	−1.324	−21.891	−21.887	2
2	+1	+16.459	+2.988	+8.106	+8.110	5
1	−3.022	−22.155	−23.569	+5.897	+5.901	7
0	+4.681	+12.428	+27.760	+23.887	+23.889	1

TABLE 4 *Synthetic multiplication*

taken as the mean relative error of the root parameters, namely

$$\Delta p/p \approx [\Delta a_i/a_i]_{\text{mean}} \approx 5 \cdot 10^{-4}.$$

This already is in good accordance with the results of the exact error analysis given in Sec. 9 below.

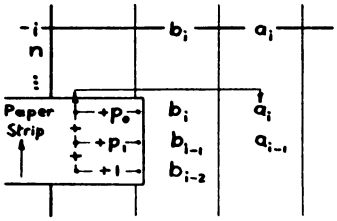

FIG. 5. a_i-Paper strip plot for synthetic multiplication

8. Variations of a_i-coefficients and p_{dh}-root parameters. Variations in the a_i-coefficients of the algebraic equation naturally produce corresponding variations in the p_{dh}-root parameters and vice versa. The first letter d of the double subscript of p_{dh} stands for the degree in the quadratic factor and is alternately 0 and 1, the second h stands for the number of the factor. For small variations there is a linear relationship

$$\Delta a_i = \sum_{dh=01}^{ih} a_{i\,dh} \Delta p_{dh} , \tag{17}$$

or by inversion

$$\Delta p_{dh} = \sum_{i=0}^{n-1} p_{dh\,i} \Delta a_i . \tag{18}$$

The coefficients a_{idh} and p_{dhi} are the matrix elements of the linear systems, and the objective is to compute them numerically. Those matrices are very useful for application, as they give the interdependence between changes of the root parameters or motion characteristics p_{dh} and the physical constants a_i of a dynamic system for example.

To deduce the matrix a_{idh} of Eq. (17), the original Eq. (1) is set up with the hth quadratic factor

$$\sum_{i=0}^{n} a_i z^i = \sum_{j=2}^{n} b_{jh} z^{j-2}(z^2 + p_{1h} z + p_{0h}). \tag{19}$$

The indices i and j have here to be discriminated because they start from different origins. By differentiation of a_i and p_{0h} term by term

$$\Delta a_i z^i = b_{jh} z^{j-2} \Delta p_{0h} ,$$

after shortening with $z^i = z^{j-2}$ and comparing with Eq. (17)

$$a_{i0h} = b_{jh} , \qquad i = j - 2 = 0 \cdots n - 2. \tag{20}$$

Differentiating a_i and p_{ih} term by term

$$\Delta a_i z^i = b_{jh} z^{j-1} \Delta p_{1h} ,$$

after shortening with $z^i = z^{j-1}$ and comparing with Eq. (17)

$$a_{i1h} = b_{jh} , \qquad i = j - 1 = 1 \cdots n - 1. \tag{20'}$$

Thus the matrix elements a_{idh} are identical with the b_{ih}-coefficients of Table 3, only lowered in their degree by 2 and 1 respectively. For completeness the a_{idh}-matrix is compiled in Table 5.

The matrix elements $p_{dh\,i}$ of Eq. (18) can now be derived by simply inversing the matrix of Eq. (17) according to the well-known algorithm of Gauss-Cholesky-Banachiewicz [5] and [8]. This has actually been done for this example with sufficient accuracy.

dh	01	11	02	12	03	13	04	14
i $a_{i\,dh}$								
7	0	+1	0	+1	0	+1	0	+1
6	+1	-3.955	+1	-5.098	+1	+0.007	+1	+0.010
5	-3.955	+6.094	-5.098	+11.623	+0.007	+0.591	+0.010	-1.426
4	+6.094	-1.322	+11.623	-11.831	+0.591	+2.786	-1.426	-3.335
3	-1.322	+2.989	-11.831	+5.697	+2.786	+13.827	-3.335	+3.583
2	+2.989	-23.568	+5.697	-7.341	+13.827	+12.454	+3.583	+4.553
1	-23.568	+27.763	-7.341	+10.694	+12.454	+8.993	+4.553	+5.096
0	+27.763	0	+10.694	0	+8.993	0	+5.096	0

TABLE 5 *Matrix* $a_{i\,dh}$

But the labor for solving this 8th order system is already rather high; it could be alleviated eventually by using automatic calculating machines, suitable for this type of problem, like that by R. R. M. Mallock [9].

However, there is another and direct approach to this problem working even more quickly and accurately. The coefficient a_i of Eq. (1) is given an increment $\Delta a_i = 1$. At first its effect $b_{1h\,i}$, $b_{0h\,i}$ on the remainder vector for the specific root factor h is computed according to the reduction formulae (4), (5) and the paper strip plot of Fig. 2;

i	a_i	b_i
n	0	0
⋮		
i	$a_i = 1$	$b_i = 1$
⋮		
k	0	$b_k = -p_0 b_{k+2} - p_1 b_{k+1}$
⋮		
1	0	$b_1 = -p_0 b_3 - p_1 b_2 \equiv b_{1h\,i}$
0	0	$b_0 = -p_0 b_2 \equiv b_{0h\,i}$

TABLE 6 *Singular formation of increment vector* $b_{dh\,i}$

this leads to the formation of the increment vector $b_{dh\,i}$ for the order i (Table 6). Now the values $b_{dh\,i}$ are built up step by step beginning with $i = 0$, and one sees that they can be telescoped into one compound for $i = 0$ to $n - 1$, as in Table 7; there is no incre-

i	$b_{1h\,i}$	$b_{0h\,i}$
0	0	+1
1	+1	0
2	$-p_{1h}\, b_{1h\,1}$	$-p_{0h}\, b_{1h\,1}$
3	$-p_{0h}\, b_{1h\,1} - p_{1h}\, b_{1h\,2}$	$-p_{0h}\, b_{1h\,2}$
⋮		
i	$-p_{0h}\, b_{1h\,i-2} - p_{1h}\, b_{1h\,i-1}$	$-p_{0h}\, b_{1h\,i-1}$
⋮		
n-1		

TABLE 7 *Telescoped array of increment vectors* $b_{dh\,i}$

ment for a_n, because it is assumed that $a_n = 1$ and constant. The computation is conveniently made along a paper strip plot, as in Fig. 6. Thereafter the increment vector

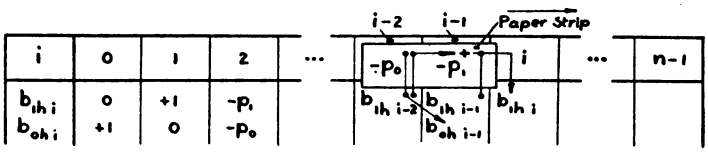

FIG. 6. Paper strip plot for $b_{dh\,i}$-increment vector

$b_{dh\,i}$ has simply to be transformed into the increment vector $p_{dh\,i}$ by means of the transform matrix K, already compiled in Table 3 for the original a_i

$$p_{dh\,i} = K \cdot b_{dh\,i}\,, \tag{21}$$

symbolizing the matrix multiplication of Eq. (7). The numerical computations are compiled in Table 8.

h K	p_{oh} / p_{ih}	$b_{dh\,i}$ / $p_{dh\,i}$	i → 0	1	2	3	4	5	6	7	$\Sigma b^2_{dh\,i}$ / $\Sigma p^2_{dh\,i}$	$b_{dh\,m}$ / $p_{dh\,m}$
1	+0.8605	d 1	0	+1	-0.943	+0.028	+0.785	-0.764	+0.045	+0.615	3.488	1.87
	+0.9430	0	+1	0	-0.861	+0.811	-0.024	-0.675	+0.657	-0.039	3.288	1.81
-.0101	+.0288	0	+0.029	-0.010	-0.016	+0.023	-0.008	-0.012	+0.019	-0.007	.00234	0.049
+.0187	+.0103	1	+0.010	+0.019	-0.027	+0.009	+0.014	-0.021	+0.008	+0.011	.00209	0.046
2	+2.234	1	0	+1	-2.086	+2.117	+0.244	-3.238	+10.38	-9.955	244.2	15.6
	+2.086	0	+1	0	-2.234	+4.660	-4.729	-0.545	+11.70	-23.19	725.2	26.9
-.0017	-.0102	0	-.0103	-.0015	+0.026	-0.051	+0.048	+0.013	-0.136	+0.254	.0889	0.298
-.0114	+.0004	1	+.0003	-.0114	+0.023	-0.023	-0.004	+0.060	-0.115	+0.107	.0295	0.172
3	+2.655	1	0	+1	+3.019	+6.459	+11.49	+17.52	+22.41	+21.14	1440	38.0
	-3.019	0	+1	0	-2.655	-8.015	-17.15	-30.49	-46.53	-59.50	7002	83.7
+.0441	+.0327	0	+0.031	+0.042	+0.044	+0.022	-0.052	-0.214	-0.508	-0.967	1.247	1.12
-.0176	-.0168	1	-0.016	-0.017	-0.008	+0.021	+0.085	+0.200	+0.377	+0.609	0.561	0.75
4	+4.681	1	0	+1	+3.022	+4.451	-0.695	-22.94	-66.06	-92.27	13430	115.9
	-3.022	0	+1	0	-4.681	-14.15	-20.84	+3.253	+107.4	+309.2	107800	328.3
-.0177	-.0010	0	-0.001	-0.018	-0.049	-0.065	+0.033	+0.403	+1.062	+1.324	3.051	1.75
+.0104	+.0040	1	+0.004	+0.010	+0.013	-0.010	-0.091	-0.226	-0.258	+0.277	0.203	0.45

TABLE 8 *Increment vectors* $b_{dh\,i}$, $p_{dh\,i}$, *error factors* $b_{dh\,m}$, $p_{dh\,m}$

For completeness the above deductions have to be modified also for linear factors $(z + p)$. Therefore Eqs. (17), (18), (19) have to be supplemented respectively by the terms

$$\sum a_{ih}\Delta p_h \quad (17'), \qquad \sum p_{hi}\Delta a_i \quad (18'), \qquad \sum b_{jh}z^{j-1}(z + p). \quad (19')$$

by differentiating Eqs. (19), (19′) term by term with respect to a_i and p_h

$$\Delta a_i z^i = b_{jh}z^{j-1}\Delta p_h\,,$$

after shortening with $z^i = z^{j-1}$ and comparing with Eqs. (17), (17′)

$$a_{ih} = b_{jh}\,, \qquad i = j - 1 = 0 \cdots n - 1. \tag{20'}$$

The reduction formula for the b-increment number $b_{0h i}$ of the remainder b_0 is, in analogy to the formula of Table 7,

$$b_{0h\,i} = -p_h b_{0h\,i-1}\,, \qquad \text{with } b_{0h\,0} = +1, \qquad i = 0 \cdots n-1, \tag{22}$$

from which the p-increment number p_{hi} is derived by

$$p_{hi} = k b_{0h\,i}\,,$$

where k is the known transform coefficient of the linear root h.

9. Error analysis. The errors of the a_i-coefficients are $\Delta a_i = \pm e$. The objective is to deduce the errors Δp_{dh} of the root parameters according to Eq. (18). With regard to the error propagation this means that the terms p_{dhi} have to be summed up for current i as an RMS-value. The errors Δb_{dh} of the b-remainders are formed analogously

$$\left.\begin{aligned} \Delta b_{1h} &= \pm e b_{1h\,m}\,, \\ \Delta b_{0h} &= \pm e b_{0h\,m}\,, \end{aligned}\right\} \text{with the } b\text{-error factors} \quad \left.\begin{aligned} b_{1h\,m} &= \left[\sum_{i=0}^{n-1} b_{1h\,i}^2\right]^{1/2}, \\ b_{0h\,m} &= [\sum b_{0h\,i}^2]^{1/2}, \end{aligned}\right\} \tag{23}$$

and

$$\left.\begin{aligned} \Delta p_{0h} &= \pm e p_{0h\,m}\,, \\ \Delta p_{1h} &= \pm e p_{1h\,m}\,, \end{aligned}\right\} \text{with the } p\text{-error factors} \quad \left.\begin{aligned} p_{0h\,m} &= [\sum p_{0h\,i}^2]^{1/2}, \\ p_{1h\,m} &= [\sum p_{1h\,i}]^{1/2} \end{aligned}\right\}. \tag{24}$$

In the case that the errors e are not equal, it would be quite as easy to sum up the terms of Eqs. (23), (24) with the specific and respective weights. The data required for the numerical computation of the error factors $b_{dh\,m}$ and $p_{dh\,m}$ are already contained in Table 8; only the columns for those RMS-values have to be completed. Now Table 3 can be supplemented with the numerical values of the b- and p-errors. It is interesting to see how, in the course of the successive steps, the remainder vector b_{dh} finally falls within the range of the b_{dh}-error. From then on, no additional correction Δp_{dh} would make sense. It can be seen from Table 3 that the b_{dh}-remainders of the last steps are lying just inside the b_{dh}-error range. The p_{dh}-errors as established according to Eq. (24) show that the calculated root parameters are correct to the last given digit.

10. Multiple and near multiple roots. Multiple roots offer interesting singularities which are best understood if one studies the transition of near multiple roots to the limit of multiple roots.

After having separated from a given algebraic equation all root factors by the above normal procedures—starting from the lowest and highest root alternately—at last a residual equation remains whose root factors may lie close together. The following quartic equation is given as a numerical example to demonstrate the approach

$$z^4 + 4.316z^3 + 10.035z^2 + 11.605z + 7.230 = 0.$$

In a first step the mean root M according to Sec. 5, Eq. (14) $p_0 = a_0^{1/2} = +2.689$, $p_1 = a_3/2 = +2.158$, is taken as a probing factor. The remainder vector b_1, b_0 is computed according to the procedure outlined above and results in $b_1 = -0.000802$, $b_0 = -0.000818$. It is already smaller than the last digit of the given a_i-coefficients so that the b_i-computations have to be made to the 6th digit. If the b_1, b_0-remainder lies within the b-error area, the mean root can no longer be dissolved into separate

factors, thus constituting physically and practically a double root. To decide this question the remainder vector has been plotted in the b-chart of Fig. 7 together with the b-error area pertaining to an error $e = \pm 0.0005$ of the a_i-coefficients according to Sec. 9; it follows that

$$b_{1h\,m} = 3.09, \qquad b_{0h\,m} = 6.47,$$

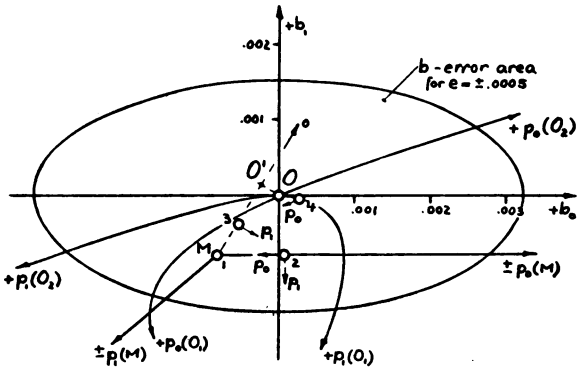

FIG. 7. b-Chart

and the $eb_{dh\,m}$ are taken as the principal axes of the error ellipse. One sees from Fig. 7 that the remainder vector b_1, b_0 of point M (step 1) actually lies already within the error area.

For further clarification the details of the b- and p-charts have to be imaged into each other reciprocally. First a series of p_0, p_1-parameters are chosen along a set of p_0- and p_1-traces computing turn by turn the b_1, b_0-vectors by means of the b_i-computation tables. This furnishes the images of the p-traces in the b-plane; their intersection with the b-error ellipse is interpolated back into the p-plane, giving there the actual

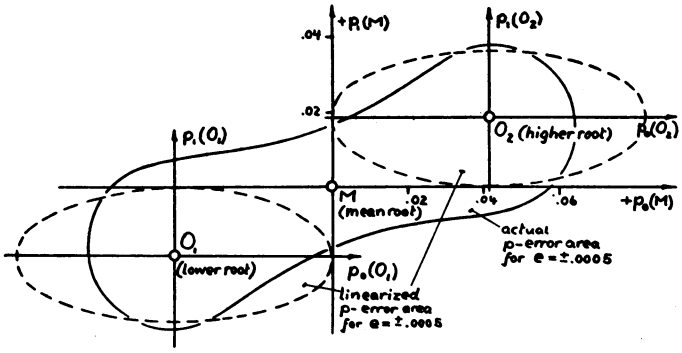

FIG. 8. p-Chart

p-error area; it is oddly shaped, but with central symmetry around the mean root point M, which obviously proves that M stands more generally for multiple roots also; it is an area overlapped several times, while the outer curve envelopes the different parts. Characteristically the p-error area is considerably enlarged compared with the b-error area by a factor of approximately 20 for this example.

The images of the p-traces in the b-plane show most impressively the singularity of point M (Fig. 7). The negative and positive portion of the $p_0(M)$- and $p_1(M)$-axes respectively fold together into straight lines touching themselves only in M with an infinitely large curvature instead of intersecting each other as in the normal case. Thus the component rays of Fig. 1 for example become indefinite, and all the procedures based on linear interdependence become inapplicable for this point M. Traces with central symmetry in the p-plane, e.g. the p_0 , p_1-axes through 0_1 , 0_2 respectively in Fig. 8, have identical images in the b-plane, but with reversed sign. The reason for this singularity is that in the relationship between the b- and p-vector in the neighborhood of M the linear terms are missing, leaving a quadratic form in p_0 , p_1

$$\left.\begin{aligned} b_1 &= c_1 + c_{1.00}p_0^2 + c_{1.01}p_0p_1 + c_{1.11}p_1^2 , \\ b_0 &= c_0 + c_{0.00}p_0^2 + c_{0.01}p_0p_1 + c_{0.11}p_1^2 , \end{aligned}\right\} \tag{25}$$

b_1 , b_0 and p_0 , p_1 denoting vector components referred to point M. If the signs of a pair of parameters p_0 , p_1 are reversed, indicating a central symmetrical p-vector, the values b_1 , b_0 (b-vector) remain unchanged, in accordance with the above Eq. (25). Since the linear terms are missing, consequently the determinant D in Eq. (12) and b in Eq. (7) tend towards zero rendering the transform matrix K indefinite and Δp in Eq. (7) indeterminate.

The next problem is to dissolve the mean root into its separate factors. But it can be solved only if the a_i-coefficients are known more accurately; it shall be assumed now that they are accurate to the 6th digit beyond the decimal point ($e = \pm 0.0000005$). Again the b_i-computation procedure is applied. But the shortcut of Sec. 5 has now to be modified for the transition from the singular point M to the linear range of the root point 0 (Fig. 7). As stated above, the linear terms in the neighborhood of M have vanished. Eq. (25) is simplified for a first rough estimate

$$b_M = cp_M^2 \qquad \text{or} \qquad p_M = Cb_M^{1/2}, \tag{26}$$

where b_M , p_M are the small vectors from M in the b- and p-plane respectively. Considering transitorily for the lower root $h = 1$ also a step o with $p_0 = p_1 = 0$, $b_1 = a_1$, $b_0 = a_0$ and interpolating the vector $b(M\text{o})$ linearly by the perpendicular $0'$ from 0 (see Fig. 7), one gets a first correction $p(M2)$ for the lower root by applying Eq. (26)

$$p(M2)/p(M\text{o}) = [b(M0')/b(M\text{o})]^{1/2}. \tag{26'}$$

As the next step 2 is planned along a p_0-trace, one takes the p_0-components instead of the p-vectors as a rough estimate; numerically

$$p(M\text{o}) = -p_0^{(1)} = -2.689, \qquad p(M2) = \Delta p_0^{(2)} = -0.03,$$

where the superscripts (1) and (2) stand for step 1 and 2 respectively (see Fig. 7). The next steps 3 and 4 are taken in a differential form along a p_1- and a p_0-trace respectively, whereby their directions are estimated according to Sec. 4 and the step increments are derived by logarithmic differentiation of Eq. (26)

$$\Delta p/p_M = \Delta b/2b_M . \tag{27}$$

A view of Fig. 7 and a check of the convergence number β show that the preliminary steps move steadily towards the root point 0. Now the transform matrix K_A can be

computed from the last three preliminary steps ($h = 1$). As a consequence of the excessive curvature around M, the convergence rate is rather slow, and one has to improve K for two steps more. The computation is concluded with step 8, where the b-remainder drops beyond its error range.

The parameters of the higher root $h = 2$ are identical with the b_i-coefficients of the last step for $h = 1$. But as a check these root parameters can also be set into their own b_i-computation table, and it follows that the b-remainder is again zero within the calculated error limits. The resulting root parameters and K-matrices are compiled in Table 9. Finally those matrices are used to compute the $p_{dh\,i}$-increment vectors and the $p_{dh\,m}$-error factors. The errors Δp for $e = \pm 0.0000005$ are included in Table 9.

h	1		2	
error Δa $\leq 10^{-7}$		error Δp		
p_{0h}	+2.64757	±.00004	+2.73081	±.00004
p_{1h}	+2.13957	±.00002	+2.17643	±.00002
$K = \begin{matrix} k_{01} & k_{00} \\ k_{11} & k_{10} \end{matrix}$	+26.6 / +22.1	+0.5 / −9.5	−26.7 / −22.1	−0.1 / +9.6

TABLE 9 *Multiple root characteristics p_{dh} and K_h*

For the transform matrix K there exists the following remarkable relation within the computational error limits, as verified by Table 9

$$K_2 = -K_1 , \tag{28}$$

meaning that all elements of the matrix reverse their sign. This theorem can be exactly proven by differentiating Eq. (25)

$$\left.\begin{aligned} \Delta b_1 &= (c_{1.00}2p_0 + c_{1.01}p_1)\Delta p_0 + (c_{1.01}p_0 + c_{1.11}2p_1)\Delta p_1 , \\ \Delta b_0 &= (c_{0.00}2p_0 + c_{0.01}p_1)\Delta p_0 + (c_{0.01}p_0 + c_{0.11}2p_1)\Delta p_1 . \end{aligned}\right\}$$

The parameters p_{0h} , p_{1h} for the root values $h = 1$ and 2 (referred to M) differ only in sign $p_{02} = -p_{01}$, $p_{12} = -p_{11}$. Consequently the matrix elements of the above equation set also reverse sign for the root values $h = 1$ and 2, and this property still holds true for the inversion matrix K, *q.e.d.* The analogous relation obviously exists also for multiple linear factors between their transform coefficients k

$$k_2 = -k_1 . \tag{28'}$$

The above K-relations can be used either as computation check or as a shortcut for the computation procedure.

At last Fig. 8 is supplemented by the p-error ellipses for the original $e = \pm 0.0005$ of the a_i-coefficients, in order to compare them with the actual p-error area. Although the linear range is considerably exceeded thereby, the linear error areas still roughly cover the actual area; they will naturally flow into each other as the matrix K tends at M to infinity.

The transition of near multiple into multiple roots can now be clearly seen in Fig. 8; the discreet roots 0_1 , 0_2 move simply towards M deforming gradually their error area, until it reaches finally a double symmetric shape with reference to the $p_0(M)$- and $p_1(M)$-axes. Physically and practically there is no difference between multiple and near multiple roots, as their large p-error area covers their small variations.

References

1. H. C. Hart and J. Travis, *Mechanical solution of algebraic equations*, J. Franklin Institute **225**, 63 (1938)
2. W. R. Evans, *Control system synthesis by root locus method*, Elec. Eng. **69**, 405 (1950)
3. F. W. J. Olver, *The evaluation of zeros of high degree polynomials*, Phil. Trans. Roy. Soc. **244**, 385 (1952)
4. L. Bairstow, Advisory Committee for Aeronautics, Reports and Memoranda No. 154, 51 (1914)
5. W. E. Milne, *Numerical calculus*, Princeton University Press, 1949
6. S. N. Lin, *A method of successive approximations of evaluating the real and complex roots of cubic and higher order equations*, J. Math. and Phys. **20**, 231 (1941)
7. F. B. Hildebrand, *Note on S. H. Lin's method of factoring polynomials*, J. Math. and Phys. **32**, 164 (1953)
8. R. Zurmühl, Matrizen, Springer Verlag, Berlin, 1950
9. R. R. M. Mallock, *An electrical calculating machine*, Proc. Roy. Soc. (London) **140A**, 457 (1933)