Linear and Multilinear Alg. [illegible]981, Vol. 10, pp. 26[illegible]-[illegible]08
0308-1087/81/1004-0265 $ [illegible]0/0

The Method of Symetric and Hermitian Forms i[illegible]ory of [illegible] Separation of the Roots of Algebraic Equations

M. G. KREIN and M. A. NAIMARK
Research Institute for Mathematics and Mechanics at Kharkov State University

Originally published in Kharkov (1936)

Translation prepared by O. BOSHKO and J. L. HOWLAND
Department of Mathematics, University of Ottawa, Ottawa, Ontario, Canada

(*Received February 27, 1981*)

Contents†

Preface 266 (3)
§ 1. The Hermite–Jacobi Method 266 (5)
  1. The origin of the method 266 (5)
  2. The separation of real roots 269 (7)
  3. The separation of the solutions of a system of equations 273 (11)
§ 2. The Fundamental Properties and Applications of Bezoutiants 275 (13)
  1. The Bezout matrix 275 (13)
  2. Signature of the Bezoutiant and index of the fraction g/f 280 (17)
  3. Determination of the index in a finite interval 283 (20)
  4. The association of the Bezoutiant with the Sturm theorem 285 (22)
  5. A transformation of the Bezoutiant 288 (24)

† Publisher's note: the page numbers give the page of the journal. The numbers in brackets are the page numbers of the original Russian version. Page numbers referred to in the text are those of the Russian version. Please refer to the *Comment by the Editors* at the end of the article.

§ 3. The Hermite Method for the Separation of Complex Roots and its Development 288 (24)
1. Hermite's results 288 (24)
2. The Hurwitz problem 292 (28)
3. The Schur–Cohn problem and the theory of symmetrical polynomials 294 (29)
§ 4. Association with Some Problems of Function Theory 300 (34)
Note by N. I. Achiezer with respect to his paper. "On a 'minimum' problem of function theory and on the number of roots of an algebraic equation which are inside the unit circle." 303 (37)
Index of the literature 304 (39)
Supplementary list of the literature relevant to § 4 307 (42)
List of text–books 308 (42)

—3—

Preface

The present publication is in the nature of a survey, therefore the presentation is similar to articles in the mathematical encyclopediae.

The authors did not intend to give a systematic presentation of all relevant problems on the basis of some single method, but rather have attempted to include different methods of establishing the various propositions, indicating their characteristic differences.

All basic methods and results are presented in detail; moreover, the indications given in the paper will allow a more or less experienced mathematician,to construct proofs by himself of nearly all the results given.

The Authors.

—5—

§ 1. THE HERMITE–JACOBI METHOD

1. The origin of the method

The first application of the law of inertia of quadratic forms to the investigation of the roots of algebraic equations is found in the proof which Jacobi [5] supplied to the well-known theorem by Borchardt [4].†

I. *The number of different pairs of complex-conjugate roots of the real polynomial*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots a_n = a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n)$$

288 (24)
288 (24)
292 (28)

294 (29)
300 (34)
n'
ic
303 (37)
304 (39)
307 (42)
308 (42)

esentation

ıll relevant
empted to
indicating

eover, the
xperienced
ılts given.

ıe Authors.

ms to the
roof which
.†

ɔf the real

α_n)

is equal to the number of variations in sign in the sequence of determinants

$$\Delta_n = \begin{vmatrix} s_0 & s_1 & \cdots & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & \cdots & s_n \\ \cdot & \cdot & \cdots & \cdots & \cdot \\ s_{n-1} & s_n & \cdots & \cdots & s_{2n-2} \end{vmatrix} \ldots \Delta_2 = \begin{vmatrix} s_0 & s_1 \\ s_1 & s_2 \end{vmatrix} \Delta_1 = s_0;\ \Delta_0 = 1$$

where s_k*–Newton sum:* $s_k = \alpha_1^k + \alpha_2^k + \ldots \alpha_n^k (k = 0, 1, 2, \ldots)$.

In this theorem Borchardt has assumed that there are no zeros in the sequence $\Delta_n \ldots \Delta_1$, Δ_0 of numbers. He obtained this theorem, taking as basis Sturm's [68] theorem and Sylvester's analysis of the Sturm functions.

Jacobi, using the theory of quadratic forms, arrived at more detailed results. His method consisted in the following:

Consider the quadratic form

$$J = \sum_{k=1}^{n} (x_0 + \alpha_k x_1 + \alpha_k^2 x_2 + \cdots + \alpha_k^{n-1} x_{n-1})^2$$

It is easy to see that

$$J = \sum_{k,l=0}^{n-1} s_{k+l} x_k x_l$$

Each pair of squares, corresponding to two complex-conjugate roots, gives sums of the form:

$$(P + Qi)^2 + (P - Qi)^2 = 2P^2 - 2Q^2$$

where P and Q are real, linear functions of the variables x_0, x_1, ..., x_{n-1}. Thus, if among the roots there are p different real

† This proof was published by Borchardt after Jacobi's death.

and q different pairs of complex-conjugate roots, then the number of positive squares of the form J equals $p+q$, and the number of negative squares equals q.

Therefore, the following theorem is valid:

II. *If* π *is the number of positive, and* ν *is the number of negative squares of the form* J*, then the polynomial* $f(x)$ *has* ν *different pairs of complex-conjugate roots and* $\pi - \nu$ *different real roots.*

On the other hand, he (Jacobi [35]) has established the following identity:

$$F = \sum_{k,l=0}^{n-1} a_{kl} x_k x_l = \frac{X_0^2}{D_0 D_1} + \frac{X_1^2}{D_1 D_2} + \cdots + \frac{X_{r-1}^2}{D_{r-1} D_r};$$

where

$$D_0 = 1;\ D_1 = a_{00},\ D_2 = \begin{vmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{vmatrix} \dots D_r = \begin{vmatrix} a_{00} & \cdots & a_{0r-1} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{r-10} & \cdots & a_{r-1r-1} \end{vmatrix} \qquad (1)$$

r = rank of the form F, and

$$X_{k-1} = \begin{vmatrix} a_{00} & a_{01} & \cdots & a_{0k-1} & \dfrac{1}{2}\dfrac{\partial F}{\partial x_0} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_{k-1,0} & a_{k-1,1} & \cdots & a_{k-1,k-1} & \dfrac{1}{2}\dfrac{\partial F}{\partial x_{k-1}} \end{vmatrix} \quad (k = 1, 2, \dots, r);$$

where it is assumed that $D_k \neq 0$ $(k = 0, 1, \dots, r)$. It follows from this identity that the number of negative squares of the quadratic form F equals the number of alternations in sign in the sequence (1).

Taking, as basis, this theorem I it follows from theorem II, and it is sufficient to assume, that there are no two consecutive zeros in this sequence.†

Independently from Jacobi, Hermite [29], [30], [31] arrived at the method of forms and developed most extensively this method and found important applications. Taking as basis the law of inertia of quadratic forms, discovered by Sylvester [73] in 1852 and, before Sylvester,

† The point is, that Gundelfinger [22] generalized the Jacobi rule for the case in which, in this sequence, two consecutive zeros are absent, and he proved that on both sides of the zero-term are terms with different signs, so that any sign could be ascribed to the zero-term. Note, moreover, that the analysis by Frobenius [16] of the Bezout forms (see pp. 21, 22).

always gives the means of determining from the entire sequence of major minors $\Delta_0, \Delta_1, \Delta_2 \dots \Delta_n \dots$ the numbers π and ν.

By the way, $r = \pi + \nu$ equals the number of the last of the minors, which is different from zero for any m:

$$J = \sum_{k,l=0}^{n-1} s_{k+l} x_n x_l = \sum_{j=1}^{m} (x_0 + \alpha_j x_1 + \dots \alpha_j^{m-1} x_{m-1})^2$$

(For the application of this rule see Baur [2].)

by Jacobi [36], and generalizing it over the Hermitian forms, Hermite arrives at rules for the separation of both real and complex roots of algebraic equations purely algebraically, and generalizes Sturm's theorem over systems of equations. He also succeeds with the establishment, in a purely algebraic way, of separate results from the index theory of Cauchy. He gives rules for the determination of the number of roots in a rectangle, circle and in an endless multitude of other algebraic curves. All analyses by this brilliant algebraist have a certain methodological significance formulated by him in the letter to Borchardt, (see [31], p. 39) "La théorie des formes quadratiques vient ainsi donner pour ces théorèmes des démonstrations indépendants de toute considérations de continuité..." (translated as: "The theory of quadratic forms thus gives the means of proving all these theorems independently of any considerations of continuity....")

2. The separation of real roots

Let us present the Hermite method, in its application to the problem of the separation of real roots,† in a slightly generalized way, as has been done by many later authors. Let

$$\Psi_k(x) = b_0^{(k)}x^{n-1} + b_1^{(k)}x^{n-2} + \cdots + b_{n-1}^{(k)} \; (k = 0, 1, \ldots, (n-1))$$

be linearly independent real polynomials, so that the determinant

$$|b_j^{(k)}|_0^{n-1} \neq 0.$$

To each equation

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n = 0,$$

with real coefficients, with roots $\alpha_1, \alpha_2, \ldots, \alpha_n$, we ascribe a quadratic form in the variables $x_0, x_1, \ldots, x_{n-1}$:

$$H = \sum_{j=1}^{n} \chi(\alpha_j)[x_0\Psi_0(\alpha_j) + x_1\Psi_1(\alpha_j) + \cdots + x_{n-1}\Psi_{n-1}(\alpha_j)]^2$$

$$= \sum_{k,l=0}^{n-1} c_{kl}x_kx_l \left[c_{kl} = \sum_{j=1}^{n} \chi(\alpha_j)\Psi_k(a_j)\Psi_l(a_j)\right].$$

where $\chi(x)$ is an arbitrarily chosen rational function which is not zero at any of the points α_j. The coefficients c_{kl}, as symmetric functions of the α_j, belong to the same field as the coefficients of the functions Ψ_j, f and χ. It is easy to obtain the formula:

$$\Delta_k = \Sigma\chi(\alpha_{j_1})\chi(\alpha_{j_2})\ldots\chi(\alpha_{j_k}) \begin{vmatrix} \Psi_0(\alpha_{j_1}) & \Psi_1(\alpha_{j_1}) & \cdots & \Psi_{k-1}(\alpha_{j_1}) \\ \cdot & \cdot & \cdots & \cdot \\ \Psi_0(\alpha_{j_k}) & \Psi_1(\alpha_{j_k}) & \cdots & \Psi_{k-1}(\alpha_{j_k}) \end{vmatrix}^2 \quad (2)$$

† In this question, as Kronecker [43] observes quite justly, one should place near the name of Hermite the name of Sylvester, who, using only one insignificant hint given him by Hermite (see Sylvester [74], p. 483) has by himself developed and generalized this method quite independently. It is even possible that the idea of the application of forms to the theory of algebraic equations occurred to Hermite himself not only in connection with his analysis of arithmetics over the forms, but also in connection with Sylvester's [72] representation of the Sturm functions.

— 8 —

for the kth principal minor $\Delta_k = [c_{ij}]_0^{k-1}$ of this form, where the summation is extended over all possible combinations of k values $j_1, j_2, \ldots, j_k$ from $1, 2, \ldots, n$.

If $\beta_1, \beta_2, \ldots, \beta_p$ are all the different real roots of the equation $f(x) = 0$ with respective multiplicities $h_1 \ldots h_p$, and if $\gamma_1, \bar{\gamma}_1; \gamma_2, \bar{\gamma}_2; \ldots; \gamma_q, \bar{\gamma}_q$ are all its pairs of different complex-conjugate roots with respective multiplicities $g_1 \ldots g_q$, then, assuming:

$$R_k = x_0\Psi_0(\beta_k) + \cdots + x_{n-1}\Psi_{n-1}(\beta_k) \quad (k = 1, 2, \ldots, p)$$

$$P_k + iQ_k = \sqrt{\chi(\gamma_k)}[x_0\Psi_0(\gamma_k) + \cdots + x_{n-1}\Psi_{n-1}(\gamma_k)] \quad (k = 1, 2, \ldots, q)$$

we obtain

$$H = \sum_{k=1}^{p} h_k\chi(\beta_k)R_k^2 + \sum_{k=1}^{q} 2g_k[P_k^2 - Q_k^2]$$

Here, as the matrix $\|b_j^{(k)}\|_0^{n-1}$ is non-singular, the rank of the matrix

$$\|\Psi_k(\alpha_j)\| = \|b_j^{(k)}\| \cdot \|\alpha_j^{n-k-1}\| \dagger$$

equals the rank of the matrix $\|\alpha_j^{k-1}\|$ i.e. equals $p+2q$. Consequently, the rank r of the form H also equals $p+2q$. Hence, since the forms R_k, P_k, Q_k are linearly independent, the following theorem results:

III. Theorem of Hermite [30]-Sylvester [74]. *If π is the number of positive and ν is the number of negative squares of the form H, and q is the number of different pairs of complex-conjugate roots of the equation $f(x) = 0$, then the latter has $\pi - q$ different real roots for which $\chi(x) > 0$ and $\nu - q$ different real roots for which $\chi(x) < 0$.*

Let us now give the most interesting results which can be obtained by selecting the functions χ, Ψ_k $(k = 0, 1, \ldots, n-1)$ in different ways:

a) Suppose that $\chi(x) = x^{2\mu}$, $\Psi_k(x) = x$ $(k = 0, 1, \ldots, n-1)$ then we obtain

$$H = \sum_{k,l=0}^{n-1} s_{2\mu+k+l}x_kx_l.$$

This form can replace the one in Theorem II, then this theorem can be generalized to some extent.

b) Assume that $\chi(x) = t - x$, $\Psi_k(x) = x^k$ $(k = 0, 1, \ldots, n-1)$ then we obtain

$$H(t) = \sum_{k,l=0}^{n-1} (s_{k+l}t - s_{k+l+1})x_kx_l.$$

It follows from Theorem III that the number of different real roots of the polynomial $f(x)$—i.e. roots that are in the interval (a, b) [on the assumption that a and b do not coincide with any of the roots of $f(x)$] equals the difference of the numbers of negative squares of the forms $H(a)$ and $H(b)$.

Wherefrom we arrive at the rule, obtained by Joachimstal [39] in a different way:

---

† The superscript indicates the row number.

— 9 —

IV. *The number of different real roots in the interval* (a, b) *of the polynomial* $f(x)$ *equals the loss in the number of variations in sign when proceeding from a to b in the series of polynomials*

$$1, \delta_1(t), \delta_2(t), \ldots, \delta_r(t)$$

where [*see* (2)]

$$\Delta_k(t) = \begin{vmatrix} s_0 & t-s_1 & s_1t-s_2 & \cdots & s_{k-1}t-s_k \\ s_1 & t-s_2 & s_2t-s_3 & \cdots & s_kt-s_{k+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ s_{k-1} & t-s_k & s_kt-s_{k+1} & \cdots & s_{2k-2}t-s_{2k-1} \end{vmatrix} = \begin{vmatrix} s_0 & s_1 & \cdots & s_k \\ s_1 & s_2 & \cdots & s_{k+1} \\ \cdot & \cdot & \cdots & \cdot \\ s_{k-1} & s_k & \cdots & s_{2k-1} \\ 1 & t & \cdots & t^k \end{vmatrix}$$

$$= \Sigma(t-\alpha_{j_1})(t-\alpha_{j_2})\ldots(t-\alpha_{j_k}) \begin{vmatrix} 1 & \alpha_{j_1} & \cdots & \alpha_{j_1}^{k-1} \\ 1 & \alpha_{j_2} & \cdots & \alpha_{j_2}^{k-1} \\ \cdot & \cdot & \cdots & \cdot \\ 1 & \alpha_{j_k} & \cdots & \alpha_{j_k}^{k-1} \end{vmatrix}^2 \quad (k = 1, 2, \ldots, r)$$

if, in this series, there are no two consecutive zeros at $t = a$ *or* $t = b$.

This theorem may also be extended to the more general case of the series of functions $1, \delta'_1(t), \ldots, \delta'_n(t)$ which is obtained when $\chi(x) = (t-x)\,\phi(x)$, where $\phi(x)$ is any rational function taking positive values at $x = \alpha_j$; then to s_k will correspond the generalized sums $s'_k = \sum_1^n \phi(\alpha_j)\alpha_j^k$. Joachimstal, studying the series $1, \delta'_1, \delta'_2 \ldots$ limited himself to the case of simple roots ($r = n$) and assumed that all the determinants $|s'_{k+l}|_0^{p-1}$ ($p = 0, 1, 2, \ldots, n$) differ from zero. As he has shown, with these assumptions, the functions $\delta'_k(t)$ will form common Sturm sequences with linear quotients, from which Theorem IV will result. All these limitations can be removed (see p. 20).

c) Take $\chi(x) = \dfrac{1}{t-x}$; $\Psi_k(x) = \phi(x)^k$ $(k = 0, 1, \ldots, n-1)$ where $\phi(x)$ is such a real polynomial that the remainders on division of $\phi(x)^k$ by $f(x)$ form a system of linearly independent functions. In this case, according to (2),

$$\Delta_k(t) = \sum_{j_1 \ldots j_k} \frac{W^2[\phi(\alpha_{j_1}) \ldots \phi(a_{j_k})]}{(t-\alpha_{j_1}) \ldots (t-\alpha_{j_k})},$$

where W is

$$W(z_1, z_2, \ldots, z_k) = \begin{vmatrix} 1 & z_1 & \cdots & z_1^{k-1} \\ 1 & z_2 & \cdots & z_2^{k-1} \\ \cdot & \cdot & \cdots & \cdot \\ 1 & z_k & \cdots & z_k^{k-1} \end{vmatrix}$$

— 10 —

In Theorem IV Joachimstal's polynomials $\delta_k(t)$ can be replaced by the polynomials $\Delta_k(t)$. Thus, this system of polynomials, similar to the system 1, $\delta_1(t) \ldots \delta_r(t)$ of polynomials, behaves analogously to the Sturm sequence. Hermite [30] was the first to construct this system.

d) It is possible to obtain sequences of Sturm functions by a suitable selection of the functions χ and Ψ to within positive multiplicative factors and, consequently, to prove the Sturm theorem by purely algebraic means (see Sylvester [74], p. 485).

Assume, for simplicity, that the polynomial $f(x)$ has no multiple roots. Suppose:

$$\chi(x) = \frac{f_1(x)}{f(x)} \frac{1}{t-x}; \qquad \Psi_k(x) = x^k \qquad (k = 0, 1, \ldots, n-1)$$

where $f_1(x)$ has been chosen in such a way that $\dfrac{f_1(x)}{f'(x)} > 0$ for all roots of the polynomial $f(x)$. In this case the minors $\Delta_k(t)$, multiplied by $f(t)$, coincide with the so-called Sylvester functions

$$\theta_k(t) = \sum \frac{f_1(\alpha_{j_1}) f_1(\alpha_{j_2}) \ldots f_1(\alpha_{j_k})}{f'(\alpha_{j_1}) f'(\alpha_{j_2}) \ldots f'(\alpha_{j_k})} W^2(\alpha_{j_1}, \alpha_{j_2} \ldots \alpha_{j_k}) \frac{1}{(t-\alpha_{j_1})(t-\alpha_{j_2}) \ldots (t-\alpha_{j_k})}$$

which, in the case of linear quotients, (the so-called "regular case") in the Sturm algorithm differ from Sturm functions only by positive factors, as was first discovered by Sylvester [72], and proved by Sturm [71]. As a consequence of the fundamental Theorem III we obtain the result that the number of real roots of the polynomial $f(x)$ in the interval (a, b) equals the loss in the number of variations in sign in the series 1, $\theta_1(t), \theta_2(t), \ldots, \theta_n(t)$.

In the regular case, the Sturm theorem is thus proved purely algebraically, without considerations of continuity.

Hattendo
the loss of th
the sequence
the Sturm s
Sturm theo
Kronecker [
involving, h
again, for th
only for an
In § 2 a p
case will be
e) Briosc
$(t-x)^m \frac{\omega(x)}{\theta(x)}$

$\omega(x), \theta(x)$ ar
the polyno
became con

3. The se

Hermite's
sequences c
the essentia
shown—to

be two real

be their ge
quadratic f

$H = \sum$

$= \sum$

where $\chi(x$
points $(\alpha_j,$

Hattendorff [24], [25] succeeded in proving that, for the irregular case also, the loss of the number of variations in sign in the transition of t from a to b in the sequence of Sylvester functions is identical with the corresponding loss in the Sturm sequence. That completed the first purely algebraic proof of the Sturm theorem. However, this approach appeared extremely clumsy. Later, Kronecker [42], [43] gave an elegant and short proof of the Sturm theorem, involving, however, the Sylvester functions (see Kronecker [43b], p. 105) but, again, for the regular case only. For the irregular case he succeeded in the same only for an infinite interval $(-\infty, \infty)$ (see Kronecker [43a], p. 124).

In § 2 a purely algebraical proof of the Sturm theorem for the most general case will be given, based upon the analysis of the Bezoutiant (see p. 22).

e) Brioschi [7] considered the form H on the assumption that $\chi(x) = (t-x)^m \dfrac{\omega(x)}{\theta(x)}$ where m is an odd number, positive or negative,

$\omega(x)$, $\theta(x)$ are two real polynomials, having the same sign for the real roots of the polynomial $f(x)$. With that the problems discussed in sections b) and c) became combined. The possibility of such generalization was evident by itself.

3. The separation of the solutions of a system of equations

Hermite's method, with which one can construct an infinite number of sequences of functions, each equivalent to the Sturm sequence, has, moreover, the essential advantage that it can be extended—as Hermite [29] himself has shown—to the case of several equations in several unknowns. Let

$$U(x, y) = 0, \qquad V(x, y) = 0 \tag{3}$$

be two real algebraic equations of degrees n_1 and n_2 respectively; let

$$(\alpha_1, \beta_1); (\alpha_2, \beta_2); \ldots; (\alpha_n, \beta_n) \qquad (n = n_1 n_2)$$

be their general solutions. With this system of equations let us associate the quadratic form:

$$H = \sum_{j=1}^{n} \chi(\alpha_j, \beta_j)[x_0 \Psi_0(\alpha_j, \beta_j) + x_1 \Psi_1(\alpha_j, \beta_j) + \cdots + x_{n-1} \Psi_{n-1}(\alpha_j, \beta_j)]^2$$

$$= \sum_{k,l=0}^{n-1} C_{kl} x_k x_l; \qquad \left[C_{kl} = \sum_{j=1}^{n} \chi(\alpha_j, \beta_j) \Psi_k(\alpha_j, \beta_j) \Psi_l(\alpha_j, \beta_j) \right]$$

where $\chi(x, y)$ is some real rational function, different from ∞ at the points (α_j, β_j), and $\Psi_k(x, y)$ are real integral functions. The coefficients of this form can be computed from the coefficients of the functions U, V, χ and Ψ_k $(k = 0, 1, 2, \ldots, n-1)$.

Assuming that the rank of the matrix $\|\Psi_k(\alpha_j, \beta_j)\|$ is n, we obtain the result (see likewise pp. 7, 8) that the number of negative (positive) squares of the form H equals the number of pairs of complex-conjugate solutions for which $\chi(x, y) \neq 0$ plus that number of real solutions for which $\chi(x, y) < 0$ $(\chi(x, y) > 0)$. Thus, if it is known beforehand that all the solutions of the system are real, then it will be possible to find out how many solutions are inside, outside or on the algebraic curve $\chi(x, y) = 0$. Here, by the inner and outer domains of the curve are meant the sets of points for which $\chi(x, y) < 0$ and $\chi(x, y) > 0$ respectively.

Assuming that

$$\chi(x, y) = \frac{1}{(t-x)(s-y)}; \qquad \Psi_k(x, y) = \phi(x, y)^k \qquad (k = 0, 1, \ldots, n-1)$$

we obtain the Hermite theorem:

V. *If $\nu(s, t)$ denotes the number of negative squares of the form*

$$\sum_{j=1}^{n} \frac{1}{(t-\alpha_j)(s-\beta_j)} \left[\sum_{k=0}^{n-1} x_k \phi_j^k\right]^2, \qquad \phi_j = \phi(\alpha_j, \beta_j) \qquad (j = 1, 2, \ldots, n)$$

— 12 —

then the number of solutions of the system satisfying, at the same time, the inequalities $a < x < b$, $c < y < d$ is equal to

$$\tfrac{1}{2}[\nu(a, c) - \nu(a, d) - \nu(b, c) + \nu(b, d)].$$

The number $\nu(s, t)$ equals the number of variations in sign in the series of major minors

$$\Delta_k(t, s) = \sum_{(j_1 \ldots j_k)} \frac{W^2[\phi_{j_1} \ldots \phi_{j_k}]}{(t-\alpha_{j_1}) \ldots (t-\alpha_{j_k})(s-\beta_{j_1}) \ldots (s-\beta_{j_k})} \quad (k = 0, 1, \ldots, n;\ \Delta_0 = 1).$$

Without giving a proof, Hermite [29] asserts that the functions 1, $\Delta_1(t, s), \ldots, \Delta_n(t, s)$ possess, besides the property stated in Theorem V, many other properties which are similar to properties of Sturm functions.

Assuming $\chi(x, y) = (t-x)^p(s-y)^q$ with p, q arbitrary odd numbers, and leaving Ψ_k arbitrary, it is possible to obtain as functions Δ_k sequences which have the same meaning for a rectangle as the Sturm sequence, rather than trying to obtain expressions analogous to the Sylvester functions in the case of one variable.

The evaluation of the symmetric functions C_{kl} in practice is very difficult. For the case in which $\chi(x, y)$ is a polynomial, this evaluation reduces to the determination of the sums $\sum_{j=1}^{n} \alpha_j^k \beta_j^k$. In some cases one can use Jacobi's method [34] for the evaluation of the coefficients C_{kl}.

The simplest form H for the determination of the number of solutions in a rectangle can be obtained by assuming:

$$\chi(x,y) = (t-x)(s-y), \qquad \Psi_k(x,y) = y^k \qquad (k = 0, 1, \ldots, n-1).$$

In this case $C_{kl} = stS_{k+l} - sT_{k+l} - tS_{k+l+1} + T_{k+l+1}$

where
$$S_k = \sum_1^n \beta_j^k; \qquad T_k = \sum_1^n \alpha_j \beta_j^k$$

and wherefrom

$$\Delta_k(t,s) = \begin{vmatrix} tS_0 - T_0 & tS_1 - T_1 & \cdots & tS_k - T_k \\ tS_1 - T_1 & tS_2 - T_2 & \cdots & tS_{k+1} - T_{k+1} \\ . & . & \cdots & . \\ tS_{k-1} - T_{k-1}, & tS_k - T_k & \cdots & tS_{2k-1} - T_{2k-1} \\ 1, & s, & \ldots, & s^k \end{vmatrix}$$

These functions are analogous to the Joachimstal functions $\delta_k(t)$ (see p. 9). They were first introduced by Hermite [29] for a particular system of two equations, after him Brioschi [7] introduced them for the general case.

— 13 —

It is necessary to note at the end of this section that Hermite's idea, as presented here, can by no means be considered to be exhausted. The system of equations (3) has not always $n = n_1 n_2$ different solutions, there exists no general rule for the selection of the functions Ψ_k so that the matrix $\|\Psi_k(\alpha_j, \beta_j)\|$ would not be singular, etc.

§ 2. THE FUNDAMENTAL PROPERTIES AND APPLICATIONS OF BEZOUTIANTS

1. The Bezout matrix

With the polynomials

$$f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n \quad a_0 \neq 0$$
$$g(x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_n \quad b_0 \gtreqless 0$$

let us associate the form

$$B(f,g; x_0, x_1, \ldots, x_{n-1}) = -B(g,f, x_0, x_1, \ldots, x_{n-1}) = \sum_{k,l=0}^{n-1} c_{kl} x_k x_l$$

which Sylvester [74] named "Bezoutiant", and whose coefficients are computed from a_k, b_l in the following way:

$$c_{kl} = d_{0,k+l+1} + d_{1,k+l} + \cdots + d_{k,l+1} \qquad (d_{kl} = a_{n-l}b_{n-k} - a_{n-k}b_{n-l}).$$

As Cayley [11] observed, the numbers c_{kl} are the coefficients of $x^k y^l$ of the integral function $\dfrac{f(x)g(y)-f(y)g(x)}{x-y}$, that is

$$\frac{f(x)g(y)-f(y)g(x)}{x-y} = \sum_{k,l=0}^{n-1} c_{kl}x^k y^l;$$

therefore this function is named the "generating function" of the Bezoutiant.

The following propositions ensue:

VI. *The determinant of the Bezoutiant equals the resultant*† *of f and g, multiplied by $(-1)^n$, that is,*

$$|c_{kl}|_0^{n-1} = (-1)^n \cdot R_{fg} = (-1)^n \begin{vmatrix} a_0 & a_1 & \cdots & a_{n-1} & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & a_{n-2} & a_{n-1} & a_n & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_0 & a_1 & a_2 & a_3 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{n-1} & b_n & 0 & 0 & \cdots & 0 \\ 0 & b_0 & \cdots & b_{n-2} & b_{n-1} & b_n & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & b_0 & b_1 & b_2 & b_3 & \cdots & b_n \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \end{matrix}$$

VII. *The defect of the Bezoutiant equals the degree of the greatest common divisor of the polynomials f and g.*

† To use the common terminology, the determinant R_{fg} on the r.h.s. could be called resultant of the polynomials f and g only when the polynomials f and g are of the same degree. However, it is easy to show that, if the degree m of the polynomial g is less than the degree n of f, then R_{fg} differs from the common resultant by the factor a_0^{n-m} (see Frobenius [16], p. 419).

Before showing how these propositions may be established, let us associate the Bezoutiant with the form

$$S(f, g; x_0, x_1, \ldots, x_{n-1}) = -S(g, f; x_0, x_1, \ldots, x_{n-1}) = \sum_{k,l=0}^{n-1} s_{k+l}x_k x_l$$

where the numbers $s_0, s_1, \ldots, s_{2n-2}$ are determined from the expansion

$$\frac{g(x)}{f(x)} = s_{-1} + \frac{s_0}{x} + \frac{s_1}{x^2} + \frac{s_2}{x^3} + \cdots$$

and, consequently, from the recurrence formulae:

$$b_k = a_0 s_{k-1} + a_1 s_{k-2} + \cdots + a_k s_{-1} \qquad (k = 0, 1, 2, \ldots) \tag{4}$$

We obtain

$$\frac{f(x)g(y)-f(y)g(x)}{x-y} = f(x)f(y)\frac{\dfrac{g(y)}{f(y)} - \dfrac{g(x)}{f(x)}}{x-y}$$

$$= f(x)f(y) \sum_{k,l=0}^{\infty} s_{k+l} x^{-(k+1)} y^{-(l+1)}$$

$$= \sum_{k,l=0}^{\infty} s_{k+l}(a_0 x^{n-k-1} + a_1 x^{n-k-2} + \cdots + a_n x^{-k-1})(a_0 y^{n-l-1} + a_1 y^{n-l-2} + \cdots + a_n y^{-l-1})$$

$$= \sum_{k,l=0}^{n-1} s_{k+l}(a_0 x^{n-k-l} + a_1 x^{n-k-2} + \cdots + a_{n-k-1})(a_0 y^{n-l-1} + a_1 y^{n-l-2} + \cdots + a_{n-l-1})$$

wherefrom

$$S(f, g;\, u_0, u_1, \ldots, u_{n-1}) = B(f, g;\, x_0, x_1, \ldots, x_{n-1}) \tag{5}$$

with

$$\begin{cases} u_0 = a_0 x_{n-1} + a_1 x_{n-2} + \cdots + a_{n-1} x_0 \\ u_1 = \qquad a_0 x_{n-2} + \cdots + a_{n-2} x_0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ u_{n-1} = \qquad\qquad a_0 x_0 \end{cases} \tag{6}$$

As the determinant of this transformation equals $a_0^n \neq 0$, it follows from (5) that:

VIII. *The forms $S(f, g)$ and $B(f, g)$ have the same number of positive and negative squares.*

Assuming, in (5), $x_0 = x_1 = \cdots x_{n-p-1} = 0$ and, consequently, by (6), $u_{n-1} = u_{n-2} = \cdots u_p = 0$, and observing that the discriminant of the truncated form B_p thus obtained equals the discriminant of the corresponding truncated form S_p multiplied by the square of the determinant of the transfor-

mation of the remaining variables u into the variables x, we obtain:

$$R_p = \begin{vmatrix} c_{n-p,n-p} & \cdots & c_{n-p,n-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ c_{n-1,n-p} & \cdots & c_{n-1,n-1} \end{vmatrix} = a_0^{2p} |s_{k+l}|_0^{p-1} \tag{7}$$

On the other hand, Hurwitz [33] has found an extremely simple expression for the determinant $a_0^{2p}|s_{k+l}|_0^{p-1}$ in terms of the polynomials f and g, namely

$$R_p = a_0^{2p}|s_{k+l}|_0^{p-1} = \begin{vmatrix} a_0 & b_0 & 0 & 0 & & 0 & 0 \\ a_1 & b_1 & a_0 & b_0 & & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ a_{2p-1} & b_{2p-1} & a_{2p-2} & b_{2p-2} & \cdots & a_p & b_p \end{vmatrix} \tag{8}$$

($a_k = b_k = 0$ when $k > n$).

This formula is obtained from the equalities:

$$a_0^{2p+1}|s_{k+l}|_0^{p-1} = (-1)^{\frac{p(p-1)}{2}} a_0^{2p+1} \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & s_{-1} & \cdots & s_{p-2} & s_{p-1} & \cdots & s_{2p-2} \\ 0 & 0 & \cdots & s_{p-3} & s_{p-2} & \cdots & s_{2p-3} \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & s_{-1} & s_0 & \cdots & s_{p-1} \end{vmatrix}$$

$$= (-1)^p \begin{vmatrix} a_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ & & & \\ a_p & a_{p-1} & \cdots & a_0 \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 & \cdots & \cdot & \cdots & 0 \\ 0 & s_{-1} & s_0 & \cdots & \cdot & \cdots & s_{2p-2} \\ 0 & 1 & 0 & \cdots & \cdot & \cdots & 0 \\ 0 & 0 & s_{-1} & \cdots & \cdot & \cdots & s_{2p-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ 0 & 0 & \cdot & \cdots & s_{-1} & \cdots & s_{p-1} \\ 0 & 0 & 1 & \cdots & \cdot & \cdots & 0 \end{vmatrix}$$

$$= (-1)^p \begin{vmatrix} a_0 & 0 & 0 & \cdots & 0 & 0 \\ a_1 & b_0 & a_0 & \cdots & 0 & 0 \\ a_2 & b_1 & a_1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ a_{2p} & b_{2p-1} & a_{2p-1} & \cdots & b_p & a_p \end{vmatrix}$$

$$= a_0 \begin{vmatrix} a_0 & b_0 & \cdot & \cdot & \cdots & 0 & 0 \\ a_1 & b_1 & a_0 & b_0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ a_{2p-1} & b_{2p-1} & \cdot & \cdot & \cdots & a_p & b_p \end{vmatrix}$$

— 16 —

if one multiplies the determinants row-wise and takes account of the formulae (4).

When $p = n$ we obtain the proposition VI from (8). In order to prove VII, let us designate by $D(x) = d_0x^q + \cdots + d_q (d_0 = 1)$ the greatest common divisor of the functions f and g; further let $f(x) = D(x)f^0(x)$, $g(x) = D(x)g^0(x)$.

It follows from the identity

$$\sum_{k,l=0}^{n-1} c_{kl}x^ky^l = \frac{f(x)g(y)-f(y)g(x)}{x-y} = D(x)D(y)\frac{f^0(x)g^0(y)-f^0(y)g^0(x)}{x-y}$$

$$= D(x)D(y)\sum_{k,l=0}^{n-q-1} c^0_{kl}x^ky^l$$

$$= \sum_{k,l=0}^{n-q-1} c^0_{kl}(d_0x^{q+k}+d_1x^{q+k-1}+\cdots+d_qx^k)(d_0y^{q+l}+d_1y^{q+l-1} + \cdots + d_qx^l)$$

that

$$B(f, g; x_0, x_1, \ldots, x_{n-1}) = B(f^0, g^0; u_0, u_1, \ldots, u_{n-q-1})$$

when

$$\begin{aligned} u_0 &= d_qx_0+d_{q-1}x_1+\cdots+d_0x_q \\ u_1 &= \qquad d_qx_1+\cdots+d_1x_q+d_0x_{q+1} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ u_{n-q-1} &= \qquad d_qx_{n-q-1}+\cdots+d_0x_{n-1}. \end{aligned}$$

No determinant of the form $B(f^0, g^0)$, being the "resultant" of f^0 and g^0, differs from zero; hence the rank of the form $B(f, g)$, which is equal to the rank of the form $B(f^0, g^0)$, equals $n-q$. q.e.d.

Observe that

$$\frac{g^0(x)}{f^0(x)} = \frac{g(x)}{f(x)} = s_{-1} + \frac{s_0}{x} + \frac{s_1}{x^2} + \cdots$$

consequently, because of (7);

$$R_{n-q} = |c_{kl}|_{q-1}^{n-1} = a_0^{2(n-q)}|s_{k+l}|_0^{n-q-1} = \frac{a_0^{2(n-q)}}{(a_0^0)^{2(n-q)}}|c^0_{kl}|_0^{n-q-1} \neq 0.$$

Wherefrom:

IX. *The rank of the Bezoutiant equals the order of the last major minor of the matrix* $\|c_{ik}\|_0^{n-1}$ *which does not vanish if, in constructing the consecutive major minors, one starts from the lower right-hand corner.*

In deriving Theorem VII, we used the property that the resultant vanishes when, and only when, f and g have a greatest common divisor. However, this theorem can be proven without using this property of the resultant by "splitting" the Bezoutiant according to the Sturm algorithm (see p. 22, also the proof given by Haupt [26], p. 268). After this proof, the characteristic property of the resultant can be thought of as a consequence of Theorem VII.

2. Signature of the Bezoutiant and index of the fraction g/f

Let us, after Cauchy [9], denote by $J_a^b R(x)$ the index of the function $R(x)$ in the interval (a,b), that is, the difference between the number of times the real function $R(x)$ suffers discontinuity from $-\infty$ to $+\infty$ as x passes from a to b and the number of times it suffers discontinuity from $+\infty$ to $-\infty$.

If the roots α of the polynomial $f(x)$ are simple, then

$$J_a^b \frac{g(x)}{f(x)} = \sum_{a<\alpha<b} \operatorname{sgn} \frac{g(\alpha)}{f'(\alpha)}.$$

Now let us state the following important and interesting fact:

X. *The signature ρ of the Bezoutiant of the real polynomials f and g (that is, also, of the form $S(f,g)$) equals the index of the fraction g/f in the interval $(-\infty, \infty)$ i.e.* $\rho = J_{-\infty}^{+\infty} \dfrac{g(x)}{f(x)}$.†

In case the roots of $f(x)$ are simple, the proposition follows from the general Theorem III and the formula

$$\frac{f(x)g(y)-f(y)g(x)}{x-y} = \sum_{k=1}^{n} \frac{g(\alpha_k)}{f'(\alpha_k)} \frac{f(x)}{x-\alpha_k} \frac{f(y)}{y-\alpha_k} \tag{9}$$

where the r.h.s. could be replaced with $B(f,g;x_0,x_1,\ldots,x_{n-1})$ if, on the l.h.s., in each of the polynomials $f(x)/(x-\alpha_k)$ and $f(y)/(y-a_k)$ x^l and y^l are replaced with x_l $(l=0,1,2,\ldots,n-1)$.

The formula (9) may be found in Darboux [13]; this formula follows from the expansion

$$\frac{g(x)}{f(x)} = \frac{b_0}{a_0} + \sum_{k=1}^{n} \frac{g(\alpha_k)}{f'(\alpha_k)} \frac{1}{x-\alpha_k}$$

† For the case when $f(x)$ has simple roots, the property of the Bezoutiant asserted in the theorem has been found by Sylvester in 1853 (see [74], p. 511) and by Hermite [31] in 1854 (see pp. 47–48).

When $f(x)$ has multiple roots, the reasoning becomes more complicated. In this case it is more convenient to consider the form S instead of the Bezoutiant. Hurwitz [33] has found the signature of the form $S(f, g)$ by means of an extremely beautiful method. We present it:

Suppose $\Theta(z) = x_0 + x_1 z + x_2 z^2 + \cdots + x_n z^{n-1}$ then

$$\frac{g(z)}{f(z)}\Theta^2(z) = \cdots + F_{-1}(x_0, x_1, \ldots, x_{n-1})z + F_0(x_0, \ldots, x_{n-1}) + \frac{S(f, g; x_0, \ldots, x_{n-1})}{z} + \ldots, \tag{10}$$

thus, $S(f, g; x_0, x_1, \ldots, x_{n-1})$ equals the sum of the residues of the l.h.s. of (10) with respect to all the poles of the function $g(z)/f(z)$.

Now, suppose $z = \alpha$ is some pole of $g(z)/f(z)$ of order λ, and let

$$\frac{g(\alpha+t)}{f(\alpha+t)} = \frac{c_0}{t^\lambda} + \frac{c_1}{t^{\lambda-1}} + \cdots$$

$$\Theta(\alpha+t) = \Theta_0(\alpha) + \Theta_1(\alpha)t + \Theta_2(\alpha)t^2 + \ldots,$$

where $\Theta_0(\alpha)$, $\Theta_1(\alpha)$, ... are, evidently, some linear forms in the variables $x_0, x_1, \ldots, x_{n-1}$. The residue with respect to the point α will then be

$$c_{\lambda-1}\Theta_0^2 + 2c_{\lambda-2}\Theta_0\Theta_1 + \cdots 2c_0(\Theta_0\Theta_{\lambda-1} + \Theta_1\Theta_{\lambda-2} + \ldots).$$

This residue, depending on the parity of λ, can be expressed as:

$$\Theta_0\Psi_0 + \Theta_1\Psi_1 + \cdots + \Theta_{\mu-1}\Psi_{\mu-1} \qquad (\lambda = 2\mu),$$

or

$$\Theta_0\Psi_0 + \Theta_1\Psi_1 + \cdots + \Theta_{\mu-1}\Psi_{\mu-1} + c_0\Theta_\mu^2 \qquad (\lambda = 2\mu+1),$$

where $\Psi_0, \Psi_1, \ldots$ are some linear functions of $x_0, x_1, \ldots, x_{n-1}$. If α is real, then the functions Θ and Ψ are real, and the residue can be expressed in the form:

$$\tfrac{1}{4}(\Theta_0+\Psi_0)^2 - \tfrac{1}{4}(\Theta_0-\Psi_0)^2 + \cdots + \tfrac{1}{4}(\Theta_{\mu-1}+\Psi_{\mu-1})^2 - \tfrac{1}{4}(\Theta_{\mu-1}-\Psi_{\mu-1})^2 \qquad (\lambda = 2\mu),$$

or

$$\tfrac{1}{4}(\Theta_0+\Psi_0)^2 - \tfrac{1}{4}(\Theta_0-\Psi_0)^2 + \cdots + \tfrac{1}{4}(\Theta_{\mu-1}+\Psi_{\mu-1})^2 - \tfrac{1}{4}(\Theta_{\mu-1}-\Psi_{\mu-1})^2 + c\Theta_\mu^2 \qquad (\lambda = 2\mu+1).$$

In this way, if λ is even, the residue with respect to α is represented by a form whose signature is zero; if λ is odd then the signature of the form = sgn c.

Since, to each complex pole α there corresponds a complex-conjugate pole with complex-conjugate residue, it is not difficult to see that the sum of the residues for two complex-conjugate poles is represented by a real quadratic form in the variables $x_0, x_1, \ldots, x_n$ with signature equal to zero.

Thus, the signature of $S(f, g; x_0, x_1, \ldots, x_{n-1})$ equals the sum of the sgn c for all poles of odd order, that is, $= J_{-\infty}^{\infty} \dfrac{g(x)}{f(x)}$, q.e.d.

It was tacitly assumed in our proof that all forms Θ and Ψ were linearly independent; however, if that were not the case, then the form S would have rank less than the sum of the orders of all poles, which is impossible because of Theorem V.

By the way, it follows from the theorem just proved that

$$\text{signature } S(f, 1) = \begin{cases} 0 & \text{when } n = 2p \\ \operatorname{sgn} a_0 & \text{when } n = 2p+1 \end{cases} \tag{11}$$

Also from this theorem the elegant theorem of Hurwitz [33] is obtained:

XI. *In order that the roots of the polynomials $f(x)$ and $g(x)$ be all single, real and alternate, it is necessary and sufficient that the sequence of determinants*

$$\begin{vmatrix} a_0 & b_0 & 0 & 0 & \cdots & 0 & 0 \\ a_1 & b_1 & a_0 & b_0 & \cdots & 0 & 0 \\ . & . & . & . & \cdots & . & . \\ 0 & 0 & . & . & \cdots & a_0 & b_0 \end{vmatrix}, \ldots, \begin{vmatrix} a_0 & b_0 \\ a_1 & b_1 \end{vmatrix}, 1,$$

contain either alternation of signs only, or constancy of signs only.

Indeed, if f and g satisfy the conditions of the theorem, then $J_{-\infty}^{\infty} \dfrac{g}{f} = \pm n$.

If we take $g = f'$, then in the form $S(f, f') = \sum_{k,l=0}^{n-1} s_{k+l} x_k x_l$, the values s_k are Newton sums, and we again obtain the Theorem II of Jacobi, and hence the Theorem I of Borchardt.

However, as we know, the form $S(f, f')$ is equivalent to the Bezoutiant $B(f, f')$ which can be transformed into the form

$$B(f, f') = \frac{1}{n} B(\hat{f}, f') + \frac{1}{n}(n a_0 x_{n-1} + (n-1) a_1 x_{n-2} + \cdots + a_{n-1} x_0)^2,$$

where $\hat{f}(x) = nf(x) - xf'(x)$. From here the Sylvester theorem follows (see [74], p. 513).

XII. *The number of negative squares in the Bezoutiant $B(\hat{f}, f')$ equals the number of different pairs of complex-conjugate roots of the polynomial $f(x)$.*

Sylvester, when establishing this theorem, was assuming that $f(x)$ had only simple roots; one year later this theorem was found by Hermite (see [31], p. 51), using the method of the generating function for constructing $B(f, f')$. However, it is necessary to point out that Sylvester did not know at that time the association between the forms $S(f, f')$ and $B(f, f')$; while, without doubt, Hermite did know it.

3. Determination of the index in a finite interval

Suppose that the degree of g is lower than that of f. Designate by $\rho_t = \pi_t - \nu_t$ the signature of the Bezoutiant $B_t = B(f, (t-x)g)$, where t is a real parameter. After Theorem X

$$\rho_t = J_{-\infty}^{\infty} \frac{(t-x)g}{f} = J_{-\infty}^{t} \frac{g}{f} - J_{t}^{\infty} \frac{g}{f};$$

thus

$$\rho_{t_2} = J_{-\infty}^{t_1} \frac{g}{f} + J_{t_1}^{t_2} \frac{g}{f} - J_{t_2}^{\infty} \frac{g}{f} \qquad (t_1 < t_2),$$

$$\rho_{t_1} = J_{-\infty}^{t_1} \frac{g}{f} - J_{t_1}^{t_2} \frac{g}{f} - J_{t_2}^{\infty} \frac{g}{f}$$

wherefrom

$$\nu_{t_1} - \nu_{t_2} = \tfrac{1}{2}(\rho_{t_2} - \rho_{t_1}) = J_{t_1}^{t_2} \frac{g}{f}.$$

By the way, it follows from this equality, with $g = f'$, another theorem, also discovered by Hermite [31];

XIII. *The number of different real roots of the polynomial $f(x)$ in the interval (t_1, t_2) equals the loss of the number of negative squares of the form $B(f, (t-x)f')$, when t transits from t_1 to t_2.*

Because $B(f, (t-x)f')$ is equivalent to $S(f, (t-x)f')$, this Hermite theorem, (as we will see later) follows from the Joachimstal Theorem IV and vice-versa. With respect to this theorem, see also Jamamoto [37].†

Consider the form S_t, equivalent to the form B_t; because

$$\frac{(t-x)g(x)}{f(x)} = \frac{ts_0 - s_1}{x} + \frac{ts_1 - s_2}{x^2} + \dots,$$

it follows that

$$S_t = S(f, (t-x)g) = \sum_{k,l=0}^{n-1} (ts_{k+l} - s_{k+l+1}) x_k x_l.$$

It follows from this that the index $J_{t_1}^{t_2} \frac{g}{f}$ equals the loss in the number of the sign variations in the sequence of functions

† By the way, the final result of the Jamamoto note is wrong.

— 21 —

$$D_0(t) = 1,\ D_1(t) = ts_0 - s_1,\ \dots,\ D_r(t) = \begin{vmatrix} ts_0 - s_1 & \cdots & ts_{r-1} - s_r \\ . & \cdots & . \\ ts_{r-1} - s_r & \cdots & ts_{2r-2} - s_{2r-1} \end{vmatrix}$$

if, in this sequence, there are neither two consecutive zeros at $t = t_1$ nor at $t = t_2$. This proposition is a generalization of the Proposition IV, for, if $f(x)$ has all roots different, then

$$s_j = \sum_{k=1}^{n} \frac{g(\alpha_k)}{f'(\alpha_k)} \alpha_k^j;$$

but, if $g(x) = f'(x)$, and $f(x)$ has multiple or simple roots, then the s_j will be Newton sums: $s_j = \sum_{k=1}^{n} \alpha_k^j$.

It is possible to show that, if not all determinants $|s_{k+l}|_0^{p-1}$ $(p = 1, \dots, r)$ are equal to zero, then all quotients $q_1, \dots, q_2$ in the Sturm algorithm:

$$f = q_1 f_1 - f_2,\ f_1 = q_2 f_2 - f_3, \dots, f_{r-2}$$
$$= q_{r-1} f_{r-1} - f_r,\ f_{r-1} = q_r f_r,\ (f_1 = g)$$

—are linear, and vice-versa; then the functions $D_0, D_1, \dots, D_r$ differ by positive factors only (Markov A., [49], Frobenius [16], Joachimstal [39]) from the denominators of the consecutive convergents of the continued fraction:

$$\frac{f_1}{f} = \cfrac{1}{q_1 - \cfrac{1}{q_2 \cdots - \cfrac{1}{q_r}}},$$

that is, from the functions:

$$j_0(t) = 1, j_1(t) = q_1, \dots, j_r(t) = \begin{vmatrix} q_1, & -1, & 0 & \cdots & 0 \\ -1, & q_2, & -1 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & 0 & \cdots & q_r \end{vmatrix}$$

(re these functions see also p. 23). Naturally, in this case, no two consecutive functions $D_k(t)$ vanish.

If, however, an irregular case arises, then, as it is possible to see from the example (see e.g. the Loewy note to the Sturm [68] Mémoire), $f = x^4 - 1$, $g = f'$, $t = 0$, several consecutive D_k can vanish simultaneously. However, Hattendorff [25] gave a rule for this case also which allows one to determine the index of $\frac{g}{f}$ in any arbitrary interval. It is true that Hattendorff limited his proof to the case of simple roots. Kronecker [43b] extended Hattendorff's analysis, when he considered the most general case and studied in detail the structure of the functions D_k. The Kronecker-Hattendorff formula is written:

$$J_{t_1}^{t_2} \frac{g}{f} = \frac{1}{2} \sum_{k=1}^{n} \{\operatorname{sgn} D_{k-1}(t_2) D_k(t_2) - \operatorname{sgn} D_{k-1}(t_1) D_k(t_1)\}$$

— 22 —

Frobenius [16] has given a new, very interesting basis for some of the results of Hattendorff and Kronecker, in connection with his remarkable investigations on the determination of the rank and signature of Hankel forms (see also Petr [60]).

4. The association of the Bezoutiant with the Sturm theorem

It is possible to prove the Sturm theorem purely algebraically using Bezoutiants. Let us apply the Sturm theorem to the functions f and $g = f_1$.

$$f = q_1 f_1 - f_2,\ f_1 = q_2 f_2 - f_3, \ldots f_{s-2} = q_{s-1} - f_s,\ f_{s-1} = q_s f_s \tag{12}$$

It is easy to see that

$$\frac{f(x)(t-y)f_1(y) - f(y)(t-x)f_1(x)}{x-y}$$
$$= f(x)f(y)q_1(t) + f_1(x)f_1(y)(t-x)(t-y)\frac{q_1^*(x) - q_1^*(y)}{x-y}$$
$$-[f_1(x)f_2(y) + f_2(x)f_1(y)] + \frac{f_1(x)(t-y)f_2(y) - f_1(y)(t-x)f_2(x)}{x-y}$$

where

$$q_1^*(x) = \frac{q(x) - q(t)}{x-t};$$

hence

$$B(f,(t-x)f_1) = q_1(t)u_1^2 - 2u_1u_2 + B(q_1^*, 1; v_0^{(1)}, \dots, v_{n_1-2}^{(1)})$$
$$+ B(f_1, (t-x)f_2; x_0, \dots, x_{n-n_1-1})$$

where †

$$u_1 = \ulcorner f(x) \urcorner, \; u_2 = \ulcorner f_1(x) \urcorner, \; v_0^{(1)} = \ulcorner f_1(x)(t-x) \urcorner, \dots, v_{n_1-2}^{(1)}$$
$$= \ulcorner f_1(x)(t-x)x^{n_1-2} \urcorner, \qquad n_1 = \text{degree of } q_1.$$

Transforming in a similar way $B(f_1, (t-x)f_2)$, then $B[f_2, (t-x)f_3]$ a.s.o.—we shall obtain

$$B(f,(t-x)f_1) = \sum_{k=1}^{s} q_k(t)u_k^2 - 2\sum_{k=1}^{s-1} u_k u_{k+1} + \sum_{k=1}^{s} B(q_k^*, 1; v_0^{(k)}, \dots, v_{n_k-2}^{(k)}).$$

The forms u, v are linearly independent, because their number equals the rank of the Bezoutiant B_t. According to the corollary (11) of Theorem X, the number of negative squares of the third summand is independent of t, hence, $\nu_{t_1} - \nu_{t_2} = N_{t_1} - N_{t_2}$, where N_t designates the number of negative squares in the Jacobi form

$$\sum_{k=1}^{s} q_k(t)u_k^2 - 2\sum_{k=1}^{s-1} u_k u_{k+1}. \tag{13}$$

† The transition from x^k to x_k is designated by $\ulcorner \; \urcorner$.

—23—

But N_t, according to the Gundelfinger–Jacobi rule (see p. 7), equals the number of sign alternations in the sequence

$$1, q_s, \begin{vmatrix} q_{s-1} & -1 \\ -1 & q_s \end{vmatrix}, \dots, \begin{vmatrix} q_1 & -1 & \cdots & 0 & 0 \\ -1 & q_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & -1 & q_s \end{vmatrix}$$

and thus, in the Sturm sequence

$$f_s(t), f_{s-1}(t), f_{s-2}(t), \dots, f(t),$$

which differs from the first sequence only by the factor $f_s(t)$. Thus the Sturm theorem is proved for the general form into which Cauchy [9] arranged it.

XIV. *The index of the fraction* $\dfrac{f_1(t)}{f(t)}$ *in the interval* (t_1, t_2) *equals the loss of the number of sign alternations in the Sturm sequence, when* t *goes over from* t_1 *to* t_2.

When $f_1(t) = f'(t)$, $J_{t_1}^{t_2}\dfrac{f'(t)}{f(t)}$ gives the number of real roots of the polynomial $f(t)$ in the interval (t_1, t_2), and in such cases we obtain the proposition which is usually brought to mind when discussing the Sturm theorem (see e.g. Grave [21], p. 385).

It was Kronecker's idea to apply Jacobi forms, i.e. forms of the type

$$\sum_{k=1}^{n} A_k x_k^2 - 2\sum_{k=1}^{n-1} B_k x_k x_{k+1}$$

to the problems of the separation of roots; however, he arrived at them in a different way, and obtained the Sturm theorem only for the case of linear quotients (see [43b], p. 105). The proof of the Theorem XIV given in this paper belongs to the authors [41a] of this paper.

In order to calculate the number of negative squares of the form (13), it is possible to use any system of major minors; specifically, if we take major minors starting from the upper corner—then we obtain the system of functions $j_0 = 1, j_1(t), j_2(t), \ldots, j_s(t)$, which have been discussed on page 21. The fact that the sequence of functions j_k, being a common Sturm sequence for the functions j_1 and j_{s-1}, is such that the loss of the number of sign alternations in the interval (t_1, t_2) also equals $J_{t_1}^{t_2}\dfrac{f_1}{f}$, Sturm knew already, (see Sturm [71], p. 367, Sylvester [75], p. 446).

For the regular case, when $q_k(x) = m_k x + n_k$, the following identity exists

$$B(f, f_1) = m_1 \ulcorner f_1(x) \urcorner^2 + m_2 \ulcorner f_2(x) \urcorner^2 + \cdots + m_s \ulcorner f_s(x) \urcorner^2 \qquad (14)$$

because

$$\frac{f(x)f_1(y) - f(y)f_1(x)}{x-y} = m_1 f_1(x) f_1(y) + \frac{f_1(x)f_2(y) - f_1(y)f_2(x)}{x-y}$$
$$= m_1 f_1(x) f_1(y) + \cdots + m_s f_s(x) f_s(y).$$

— 24 —

Because the degrees of the polynomials $f_1(x), f_2(x), \ldots, f_s(x)$ decrease by one, the linear functions $\ulcorner f_1(x) \urcorner, \ldots, \ulcorner f_s(x) \urcorner$, are nothing more than those linear functions which are obtained when $B(f, f_1)$ is resolved into a sum of squares by means of the well-known Lagrange [44] method. From here, Darboux, assuming the roots of $f(x)$ to be simple, obtained very wittily Sylvester's formulae for the Sturm functions, as well as the representation of these functions in terms of the coefficients of the polynomials f and f_1, as found by Cayley [10].

The Sturm Theorem XIV follows also from the formula (14) for the interval $(-\infty, \infty)$ (for a regular case).

5. A transformation of the Bezoutiant

Again assuming all roots of $f(x)$ to be distinct, we can easily obtain an associated form of the Bezoutiant from the formula (see p. 17)

$$B(f, g; x_0, \ldots, x_{n-1}) = \sum_{k=1}^{n} \frac{g(\alpha_k)}{f'(\alpha_k)} \cdot \left[\frac{f(x)}{x-\alpha_k} \right]^2$$

Assuming:

$$X_k = \frac{1}{2}\frac{\partial B}{\partial x_k} = \sum_{j=1}^{n} \frac{g(\alpha_j)}{f'(\alpha_j)} \cdot \left[\frac{f(x)}{x-\alpha_k} \right] (a_0\alpha_k^{n-1} + \cdots + a_{n-1}),$$

we obtain easily (see Darboux [13])

$$B = \sum_{k=1}^{n} \frac{1}{g(\alpha_k)f'(\alpha_k)} (X_0 + \alpha_k X_1 + \cdots + \alpha_k^{n-1} X_{n-1})^2$$
$$= \sum_{j,k=0}^{n-1} s'_{j+k} X_j X_k, \text{ where } s'_k = \sum_{j=1}^{n} \frac{\alpha_j^k}{g(\alpha_j)f'(\alpha_j)}.$$

This transformation of the Bezoutiant into Hankel form was known already to Jacobi [34a]. It is possible to prove, that also for a general case when the polynomial $f(x)$ has multiple roots, the associated form for the Bezoutiant is some Hankel fotm (see Frobenius [16]). The form $\sum_{j,k=0}^{n-1} s'_{j+k} X_j X_k$ plays an important part in the detailed investigations of Kronecker [43] on Sturm functions; however, this form can be encountered even earlier in Hermite [31].

§ 3. THE HERMITE METHOD FOR THE SEPARATION OF COMPLEX ROOTS AND ITS DEVELOPMENT

1. Hermite's results

As Hermite has shown in his famous letter to Borchardt [31], the method of forms enables us to separate the complex roots of algebraic equations with complex coefficients; however, here, instead of the theory of common quadratic forms one has to apply the theory of a new, more general class of forms which have been introduced into science also by

Hermite–Hermitian forms. In this letter, at first, Hermite solves the following problem:

Given an algebraic equation

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n = 0$$

with arbitrary complex coefficients, it is required to determine how many roots this equation has in the upper half-plane.

In order to solve this problem Hermite considers the expression

$$-i\frac{f(x)\bar{f}(y) - f(y)\bar{f}(x)}{x-y} = \sum_{k,l=0}^{n-1} A_{kl}x^k y^l$$

where $\bar{f}(x) = \bar{a}_0x^n + \bar{a}_1x^{n-1} + \cdots + \bar{a}_n$ is the polynomial with complex-conjugate coefficients.†

From this generating function Hermite constructs the form

$$H(f, x_0, \ldots, x_{n-1}) = \sum_{k,l=0}^{n-1} A_{kl}x_k x_l$$

where the coefficients A_{kl}, as it is easy to see, are real; and he proves the following theorem:

XV. *If π is the number of positive, ν is the number of negative squares of the form $H(f; x_0, \ldots, x_{n-1})$, then $f(x)$ has exactly $n-\pi-\nu$ roots in common with $\bar{f}(x)$, and, moreover, there are π roots more in the upper half-plane and ν roots more in the lower.*‡

In order to prove this we will use the method, for whose origin we are obliged to Liénard and Chipart [48] and by which, as Fujiwara [18] observed, nearly all problems discussed in this paper could be combined. We stipulate that the variables $x_0, \ldots, x_{n-1}$ can also be complex, and by $H(f; x_0, \ldots, x_{n-1})$ we understand the Hermite form $\sum_{k,l=0}^{n-1} A_{kl}x_k\bar{x}_l$. Let

$$D(x) = d_0x^p + d_1x^{p-1} + \cdots + d_p$$

be the greatest common divisor of the polynomials $f(x)$ and $\bar{f}(x)$; without limiting the generality one can consider it to be real, so that $f(x) = f^0(x)D(x)$, $\bar{f}(x) = \bar{f}^0(x)D(x)$, where $f^0(x)$, $\bar{f}^0(x)$ are mutually prime.

From the identity

$$-i\frac{f(x)\bar{f}(y) - f(y)\bar{f}(x)}{x-y} = -iD(x)D(y)\frac{f^0(x)\bar{f}^0(y) - f^0(y)\bar{f}^0(x)}{x-y}$$

it follows that

$$H(f;\, x_0, \ldots, x_{n-1}) = H(f^0;\, u_0, \ldots, u_{n-p-1}),$$

† In the following complex-conjugate values will be designated by the bar.

‡ Hermite considered only the case in which f and $\bar{f}$ have no common factors and f has only simple roots.

where:

$$\begin{aligned} u_0 &= \ulcorner D(x) \urcorner = d_p x_0 + d_{p-1} x_1 + \cdots + d_0 x_p \\ u_1 &= \ulcorner xD(x) \urcorner = d_p x_1 + \cdots + d_1 x_p + d_0 x_{p+1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_{n-p-1} &= \ulcorner x^{n-p-1} D(x) \urcorner = d_p x_{n-p-1} + \cdots + d_0 x_{n-1} \end{aligned}$$

are linearly independent. Hence the forms $H(f)$ and $H(f^0)$ have the same signatures and ranks; therefore, in order to prove the theorem it is sufficient to consider only the case in which $f(x)$ and $\bar{f}(x)$ are mutually prime.

Let $f(x) = f_1(x) f_2(x)$, where f_1 and f_2 are complex or real polynomials of degrees n_1 and $n_2 (n = n_1 + n_2)$ respectively; then, replacing in the identity

$$-i\frac{f(x)\bar{f}(y) - f(y)\bar{f}(x)}{x-y} = -if_2(x)\bar{f}_2(y)\frac{f_1(x)\bar{f}_1(y) - f_1(y)\bar{f}_1(x)}{x-y}$$

$$-if_1(y)\bar{f}_1(x)\frac{f_2(x)\bar{f}_2(y) - f_2(y)\bar{f}_2(x)}{x-y}$$

x^k by x_k and y^k by $\bar{x}_k$, we will obtain:

$$H(f;\, x_0, \ldots, x_{n-1}) = H(f_1;\, u_0, \ldots, u_{n_1-1}) + H(f_2;\, v_0, \ldots, v_{n_2-1}),$$

where the linear forms†

$$u_0 = \ulcorner f_2(x) \urcorner, \ldots, u_{n_1-1} = \ulcorner x^{n_1-1} f_2(x) \urcorner;$$

$$v_0 = \ulcorner \bar{f}_1(x) \urcorner, \ldots, v_{n_2-1} = \ulcorner x^{n_2-1} \bar{f}_1(x) \urcorner$$

are linearly independent, because their determinant differs only by a factor from the resultant of the mutually prime functions $\bar{f}_1(x)$ and $f_2(x)$. As it is possible also, in turn, to "split" the forms $H(f_1)$ and $H(f_2)$; we have

$$\frac{1}{a_0\bar{a}_0} H(f;\, x_0, \ldots, x_{n-1}) = H(x-\alpha_1;\, U_1)$$

$$+ H(x-\alpha_2;\, U_2) + \cdots + H(x-\alpha_n;\, U_n) = \sum_{k=1}^{n} \frac{\alpha_k - \bar{\alpha}_k}{i} U_k \bar{U}_k,$$

where $\alpha_1, \ldots, \alpha_n$ are the roots of $f(x)$ and $U_1, \ldots, U_n$ are linearly independent linear forms in the variables $x_0, \ldots, x_{n-1}$. From this expansion of the form H into a sum of squares the Theorem XV follows immediately.

The "splitting" of the form $H(f)$ corresponding to the expansion of f in factors is the central idea which characterizes the Liénard–Chipart method. Hermite himself has proved his theorem, proceeding from the form

$$\sum_{k,l=1}^{n} \frac{i}{\alpha_k - \bar{\alpha}_l} \xi_k \bar{\xi}_l,$$

† With regard to the symbol ⌈‾⌉ see the footnote on p. 22.

the signature and rank of which could easily be determined, and then transforming it into the form H. However, this method of proof requires superfluous limitations (see the footnote ‡, p. 25).

The Hermite form can be represented as a Bezoutiant of two real polynomials. Let $f = g + ih$, where g and h are real polynomials. Let us suppose that the degree of g is not less than the degree of h, for otherwise, we would consider if instead of f. Then it follows from the identity

$$-i\frac{f(x)\bar{f}(y) - f(y)\bar{f}(x)}{x-y} = -2\frac{g(x)h(y) - g(y)h(x)}{x-y}$$

that

$$H(f; x_0, \ldots, x_{n-1}) = -2B(g, h; x_0, \ldots, x_{n-1}).$$

From here we obtain a theorem mentioned by Sturm [69], [70] already in his mémoire on the Cauchy investigations.

XVI. *The index* $J_{-\infty}^{\infty}\frac{h}{g}$ *(the degree of* g *not less than the degree of* h*) equals the difference between the number of zeros of the function* $f = g + ih$ *located in the lower and the upper half-planes.*

The method of proof belongs to Hermite [31], who limited himself to the discussion of the case in which g has only simple roots, since for this case only he had at his disposal the Theorem X, which he derived from the Jacobi formula.

It follows from this theorem:

XVII. *If all roots of* $f = g + ih$ *are in the upper half-plane, then all roots of* g *and* h *are different, real, and alternate among themselves and vice-versa.*

The first part of this theorem is known in the literature under the name

Biehler [3]–Hermite [32], because these authors gave a proof which is not connected with Theorem X (re this theorem see the works by Laguerre [45], Jentzsch [38]).

Hermite also showed [31] how, by means of his Theorem XV, it is possible to determine how many roots of the equation $f(x) = 0$ are in the domains determined by the inequalities $V(x, y) > 0$ and $V(x, y) < 0$, where $V(x, y)$ is the imaginary part of some rational function $\phi(z)(z = x+iy)$. In order to obtain that, one has to construct the equation $F(z) = 0$ whose roots are the expressions $\phi(\alpha_1), \ldots, \phi(\alpha_n)$ (see, e.g. Grave [21]). The number of positive (negative) squares of $H(F)$ equals the number of those α_k which are located in the domain $V(x, y) > 0$ $(V(x, y) < 0)$.

If, for example, to take with Hermite

$$\phi(z) = (z-\xi-i\eta)^2, \text{ then } V(x, y) = (x-\xi)(y-\eta).$$

Wherefrom, if $\pi(\xi, \eta)$ designates the number of positive squares of the respective form $H(F)$, then the number of roots of the

equation $f(x) = 0$ which lie in the rectangle $\xi_1 < x < \xi_2, \eta_1 < y < \eta_2$ equals

$$\tfrac{1}{2}[\pi(\xi_1, \eta_1) - \pi(\xi_1, \eta_2) - \pi(\xi_2, \eta_1) + \pi(\xi_2, \eta_2)].$$

In a similar way it is possible to obtain the number of roots in an infinite strip, in a circle, in a circular ring, etc.

2. The Hurwitz problem

If it is desired to determine the number of roots of a given equation $f(x) = 0$ located in the left half-plane, then it is necessary to construct the form $H(\bar{f}; x_0, \ldots, x_{n-1})$, where $\bar{f}(x) = f(ix)$.

It is easy to see that the form $H(\bar{f}; x_0, \ldots, x_{n-1})$ can be transformed into a form $R(f; \xi_0, \ldots, \xi_{n-1})$ by setting $\xi_k = i^k x_k$ $(k = 0, \ldots, n-1)$. The form R has, as its generating expression, the function

$$R(\xi, \eta) = \frac{f(\xi)\bar{f}(\eta) - \bar{f}(-\xi)f(-\eta)}{\xi+\eta} \qquad (\xi = ix, \quad \eta = iy),$$

i.e. is obtained from $R(\xi, \eta)$ by substitution of ξ^k by ξ_k, η^k by $\bar{\xi}_k$. Assume now that the polynomial $f(x)$ is real; let further

$$g(x) = a_{n-1} + a_{n-3}x + a_{n-5}x^2 + \ldots, \qquad h(x) = a_n + a_{n-2}x + \ldots,$$

consequently $f(x) = h(x^2)+xg(x^2)$. From the identity

$$\tfrac{1}{2}R(x,y) = \frac{f(x)f(y)-f(-x)f(-y)}{2(x+y)}$$
$$= xy\frac{h(x^2)g(y^2)-h(y^2)g(x^2)}{x^2-y^2} + \frac{x^2g(x^2)h(y^2)-y^2g(y^2)h(x^2)}{x^2-y^2}$$

we obtain

$$\tfrac{1}{2}R[f;\, x_0, \ldots, x_{n-1}] = B(h, g;\, x_1, x_3, x_5, \ldots)+B(xg, h;\, x_0, x_2, x_4, \ldots).$$

In order that all roots of f be located in the left half-plane it is necessary and sufficient that the form $R(f)$ be positive definite and, consequently, that the independent forms $B(h, g; x_1, x_3, x_5, \ldots)$ and $B(xg, h; x_0, x_2, \ldots)$ would be also. Then, applying the rules for the computation of minors of the Bezoutiant (see p. 15, formula (8)), we arrive at the Hurwitz [33] theorem.

XVIII. *In order that all roots of the real equation*

$$a_0x^n+a_1x^{n-1}+\cdots+a_n(a_0>0)$$

have negative real part, it is necessary and sufficient that all determinants

$$a_1, \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}, \ldots, \begin{vmatrix} a_1 & a_3 & \cdots & a_{2n-1} \\ a_0 & a_2 & \cdots & a_{2n-2} \\ 0 & a_1 & \cdots & a_{2n-3} \\ . & . & \cdots & . \\ 0 & 0 & \cdots & a_n \end{vmatrix} (a_k = 0 \text{ for } k > n)$$

be positive.

—29—

Hurwitz discovered this theorem on the suggestion of the renowned scholar-technician Stodola [67] who applied it to a problem on the theory of turbines. The method of proof given here is in principle the same as the method used by Liénard and Chipart; the Hurwitz method (or the method used by Hurwitz himself) differs only in that, instead of the forms B, the respective forms S are considered, and instead of Hermite's theorem, the Theorem X is applied directly (see p. 17, see also Fujiwara [17]).

The positiveness of the Bezoutiant is associated with a certain order of alternation of the roots of the polynomials (see the Theorem XI); using the rule of signs by Descartes (see, e.g. Grave [21], p. 352), it is also possible to obtain from the positiveness of the Bezoutiants $B(h, g)$ and $B(xg, h)$ the Liénard–Chipart [48] criterion:

XIX. *In order that all roots of the polynomial $f(x)$ lie in the left half-plane, it is necessary and sufficient that the quadratic form $B(h, g)$ be positive definite, and that all coefficients of the polynomial $h(x)$ have the same sign as a_0.*

Liénard and Chipart have established this theorem by means of many "considerations on continuity"; however, that is not necessary. Moreover, Liénard and Chipart studied the relationship between the rank and the signature of $B(h, g)$ and the structure of the set of roots of f.

In connection with some problems in the theory of vibrations the English mechanic Routh [63], [64] worked out, independently of the Hurwitz studies, a practical algorithm which allows one to determine the number of roots of a real equation in the left and in the right half-planes.

I. Schur [66] has proposed another completely original algorithm for determining whether a real polynomial has all its roots in the left half-plane or not. Not without some interest are the investigations carried out by the Italian mathematician Orlando [54], [55], [56], [57], [58], [59] (see also Liénard [47]) in connection with the Hurwitz theorem.

As Fujiwara [18] observed, it is possible to obtain the following general result from the Hermite theorem:

XX. *In order that all roots of the real polynomial $f(x)$ lie inside the angle $-\theta < \arg x < \theta$, it is necessary and sufficient that the Hermitian form, which has as its generating function the expression*

$$-i\frac{f(e^{i\theta}x)f(e^{-i\theta}y)-f(e^{i\theta}y)f(e^{-i\theta}x)}{x-y}$$

be positive definite.

3. The Schur–Cohn problem and the theory of symmetrical polynomials

It is possible to associate with each complex polynomial

$$f(z) = a_0z^n + \cdots + a_n \quad \text{a polynomial} \quad f^*(z) = z^n\bar{f}\left(\frac{1}{z}\right) = \bar{a}_nz^n + \cdots + \bar{a}_0.$$

According to M. Krein's [41] terminology the polynomial $f(z)$ is called symmetrical if $f^*(z) = f(z)$.

The greatest common divisor of the polynomials $f(z)$, $f^*(z)$ can differ from a symmetrical polynomial only by a constant factor. In order that a polynomial may be symmetrical, or would differ from symmetrical by a constant factor, it is necessary and sufficient that, to each root α having $|\alpha| \neq 1$, the polynomial has a root $\alpha^* = \frac{1}{\bar{\alpha}}$ with the same multiplicity as α.

Let us associate, as Schur [65] and Cohn [12] do, with each polynomial $f(z)$ the Hermitian form

$$\mathscr{H}(f; x_0, \ldots, x_{n-1}) = \sum_{j=1}^{n} |a_0 x_j + a_1 x_{j+1} + \cdots + a_{n-j} x_n|^2$$
$$- \sum_{j=1}^{n} |a_n x_j + a_{n-1} x_{j+1} + \cdots + a_j x_n|^2,$$

having as its generating polynomial the expression

$$\frac{f^*(x)\bar{f}^*(y) - f(x)\bar{f}(y)}{1-xy}$$

As $f^* = f_1^* f_2^*$ follows from $f = f_1 f_2$, it is possible, using the Liénard–Chipart (see Fujiwara [18]) method, to prove, completely analogously as the Hermite theorem was proved, the following proposition:

XVa. *If π is the number of positive and ν is the number of negative squares of the form $\mathscr{H}$, then the polynomial $f(z)$ has $n-\pi-\nu$ roots in common with $f^*(z)$, and besides them, π roots lying inside the unit circle ($|z| < 1$) and ν roots outside this circle.*

It follows from this theorem that, in order that all roots of $f(z)$ be inside of the unit circle, it is necessary and sufficient that the form $\mathscr{H}(f)$ be positive definite. This result has been discovered by Schur, when he carried out his beautiful investigations on bounded functions, regular inside the unit circle (see Schur [65]). Cohn [12] has obtained the general Theorem XVa for the case in which f and f^* do not have a greatest common divisor, and also he obtained it not by way of purely algebraical reasoning, for he used the Rouché [62] theorem (see for this theorem e.g., Gursa [23], p. 99). In the most general form, and using the Liénard–Chipart method, this theorem is proved by Fujiwara [18].

As the form $\mathscr{H}$ may be transformed into the Bezoutiant $\mathscr{B}$, so the form $\mathscr{H}$ may be transformed into a form constructed from two symmetrical polynomials, which serves as a Bezoutiant in these problems. Let us define, for this purpose, two symmetrical polynomials g and h by the equalities

$$f = g + ih, \qquad f^* = g - ih.$$

Then, denoting by $\mathscr{B}(g, h; x_0, \ldots, x_{n-1})$ the form with the generating polynomial

$$\mathscr{B}(x,y) = i\frac{\mathscr{g}(x)\bar{\mathscr{h}}(y)-\bar{\mathscr{g}}(y)\mathscr{h}(x)}{1-xy}$$

we obtain easily

$$\mathscr{H}(\mathscr{f};x_0,\ldots,x_{n-1}) = 2\mathscr{B}(\mathscr{g},\mathscr{h};x_0,\ldots,x_{n-1}).$$

In turn, the form $\mathscr{B}$ is transformed from the Toeplitz form

$$\mathscr{T}(\mathscr{g},\mathscr{h};u_0,\ldots,u_{n-1}) = \sum_{k,l=0}^{n-1} c_{k-l}u_k\bar{u}_l \qquad (c_{-k} = \bar{c}_k)$$

where the numbers "c" are determined from the expansion

$$-i\frac{\mathscr{h}(z)}{\mathscr{g}(z)} = c+c_1z+c_2z^2+\ldots \qquad (c_0 = c+\bar{c}),$$

by means of the transformations

$$\begin{aligned} u_0 &= g_nx_0+g_{n-1}x_1+\cdots+g_0x_{n-1}\\ u_1 &= \qquad\quad g_nx_1\quad+\cdots+g_1x_{n-1}\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ u_{n-1} &= \qquad\qquad\qquad\qquad g_nx_{n-1} \end{aligned}$$

where $\mathscr{g}(z) = g_0z^n+g_1z^{n-1}+\ldots g_n$.

As Herglotz [28] has shown, it is possible to establish a theorem analogous to Theorem X:

Xa. *The index of the fraction* $\frac{\mathscr{h}(z)}{\mathscr{g}(z)}$ *along the unit circle, when going in the positive direction, equals the signature of the form* $\mathscr{T}$.

To prove this Herglotz proceeds from the fact that the form $\mathscr{T}$ can be represented as an integral over the unit circumference as follows:

$$\mathscr{T}(\mathscr{g},\mathscr{h};x_0,\ldots,x_{n-1}) = \frac{1}{2\pi}\oint\frac{\mathscr{h}(z)}{\mathscr{g}(z)}\theta(z)\bar{\theta}\left(\frac{1}{z}\right)\frac{dz}{z},$$

where $\theta(z) = x_0+x_1z+\cdots+x_{n-1}z^{n-1}$. This formula corresponds to a certain extent to the Hurwitz formula (see p. 18):

$$S(g,h;x_0,\ldots,x_{n-1})-\frac{1}{2\pi i}\oint\frac{h(z)}{g(z)}\theta^2(z)dz = \Sigma\ \mathrm{Res}\left(\frac{g(z)}{h(z)}\theta^2(z)\right).$$

Combining the Theorems Xa and XVa, we obtain a theorem on the index for a circle:

XVIa. *The i*

positive direc
outside and i

Theorem

transformati

half-plane $I($

where $h(x)$

$+i)^n f\left(\frac{x-i}{x+i}\right)$

into $H(f)$ an

thus

we obtain th

$(x+i)^n$

from which i

with

Thus, the
correspondi

XVIa. *The index of the quotient* $\frac{\mathfrak h}{\mathfrak g}$ *along the unit circle, when going in the positive direction, equals the difference between the numbers of roots of* $\mathfrak f = \mathfrak g + i\mathfrak h$ *outside and inside the unit circle.*

— 32 —

Theorem XVIa can be obtained from Theorem XVI by means of the transformation $z = \frac{x-i}{x+i}$, transforming the unit circle $|z| < 1$ into the upper half-plane $I(x) > 0$; indeed with this transformation $\frac{\mathfrak h(z)}{\mathfrak g(z)}$ will go over into $\frac{h(x)}{g(x)}$, where $h(x)$ and $g(x)$ are real polynomials, and $f(x) = g(x)+ih(x) = (x+i)^n\mathfrak f\left(\frac{x-i}{x+i}\right)$. Moreover, it is possible to transform the forms $\mathscr{H}(\mathfrak f)$ and $\mathscr{T}(\mathfrak f)$ into $H(f)$ and $S(f)$. Indeed, assuming

$$u = \frac{x-i}{x+i}, \quad v = \frac{y+i}{y-i}, \quad f(x) = (x+i)^n \mathfrak f\left(\frac{x-i}{x+i}\right)$$

thus

$$\bar f(x) = (x+i)^n \mathfrak f^*\left(\frac{x-i}{x+i}\right)$$

we obtain the identity

$$(x+i)^{n-1}(y-i)^{n-1}\frac{\mathfrak f^*(u)\bar{\mathfrak f}^*(v)-\mathfrak f(u)\bar{\mathfrak f}(v)}{1-uv} = -i\frac{f(x)\bar f(y)-f(y)\bar f(x)}{2(x-y)}$$

from which it follows, that

$$\mathscr{H}(\mathfrak f; u_0,\dots,u_{n-1}) = \frac{1}{2}H(f; x_0,\dots,x_{n-1})$$

with

$$\begin{aligned} u_0 &= \ulcorner (x+i)^{n-1} \urcorner = x_{n-1}-(n-1)ix_{n-2}+,\dots \\ u_1 &= \ulcorner (x+i)^{n-2}(x-i) \urcorner \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ u_{n-1} &= \ulcorner (x-i)^{n-1} \urcorner . \end{aligned}$$

Thus, the Theorems Xa, XVa, XVIa can be obtained by transforming the corresponding functions and forms from the Theorems X, XV, XVI.

The fact that, to the transformation of the functions by means of the linear fraction substitution corresponds a certain transformation of the forms, Hermite (see [31], p. 51) knew already.

Theorem XI has also its counterpart, which follows from Theorem Xa:

XIa. *In order that all roots of two symmetrical polynomials, g and h be different, lie on the circumference of the unit circle and alternate, it is necessary and sufficient that the form $\mathscr{B}(g, h)$ be definite* [28], [41].

A trigonometric polynomial $G(\phi) = e^{-\frac{in\varphi}{2}} g(e^{i\varphi})$ corresponds to each symmetrical polynomial $g(z)$ of degree n, so that, depending on whether n is even or odd, $G(\phi)$ will have the form:

$$a_0 + \sum_{k=1}^{m} (a_k \cos k\phi + b_k \sin k\phi) \qquad (n = 2m) \tag{A}$$

or

$$\sum_{k=1}^{m} \left[a_k \cos\left(k - \frac{1}{2}\right)\phi + b_k \sin\left(k - \frac{1}{2}\right)\phi \right] \qquad (n = 2m-1). \tag{B}$$

Thus, with Theorem XIa it is possible to find out if the roots of two trigonometrical polynomials alternate. It is easy to see that the derivative $G'(\phi) = -ie^{-\frac{in\varphi}{2}} g_\delta(e^{i\varphi})$, where $g_\delta(z) = \frac{n}{2} g(z) - z g'(z)$ is an anti-symmetrical polynomial, i.e., $g_\delta^*(z) = -g_\delta(z)$, and hence $i g_\delta(z)$ is a symmetrical polynomial.

From the equality

$$\mathscr{B}(g, h; 1\, \alpha, \ldots, \alpha^{n-1}) = i \frac{g(\alpha)\overline{h(\alpha)} - \overline{g(\alpha)}\, h(\alpha)}{1 - \alpha\bar{\alpha}}$$

it follows that the right hand side preserves its sign if g and h have all their alternating roots among themselves on the unit circle. Wherefrom immediately the theorem [41] follows:

XXIa. *If the roots of the symmetrical polynomials $g(z)$ and $h(z)$ (of the trigonometrical polynomials $G(\phi)$ and $H(\phi)$) all lie on the circumference of the unit circle (all are real) and alternate, then the roots of the polynomials $g_\delta(z)$ and $h_\delta(z)$($G'(\phi)$ and $H'(\phi)$) also possess the same properties.*

This theorem corresponds to A. Markov's [50], [51] theorem.

XXI. *If the roots of two real polynomials $g(x)$ and $h(x)$ are all real and alternate, then, the roots of their derivatives $g'(x)$, $h'(x)$ also alternate.*

This theorem follows also from the fact that the expression

$$g'(x)h(x)-h'(x)g(x) = B(g,h;1,x,\dots,x^{n-1})$$

preserves its sign if g and h possess the required property (see Theorem X); however, it follows also, for example, as D. Grave [20] has noticed, from the Hermite–Biehler theorem, (see p. 29) for, if the roots of $f = g+ih$ are in the upper (or lower) half-plane, then the roots of $f' = g'+ih'$ will be there also.

As the index of the fraction $i\dfrac{g_\delta(z)}{g(z)}$ along the circumference of the unit circle in the positive direction equals the number of roots $g(z)$ has on the same circle, and as $\dfrac{g_\delta(z)}{g(z)} = \dfrac{s_0}{2}+s_{-1}z+s_{-2}z^2+\dots$, where s_k are the Newton sums $(k = 0, \pm 1, \pm 2,\dots; s_{-k} = \bar{s}_k)$, then according to Theorem Xa, a theorem first formulated by Herglotz [28] is true:

IIa. *If π is the number of positive, ν is the number of negative squares of the form $\mathcal{T} = \sum_{k,l=0}^{n-1} s_{k-l}\bar{x}_k x_l$, that is of the form*

$\mathcal{B}(g, ig_\delta; x_0,\dots,x_{n-1})$, then the symmetric polynomial $g(x)$ has $\pi-\nu$ different roots on the circumference of the unit-circle and ν pairs of different roots located as mirror images in this circle.

This theorem is analogous to the Borchardt theorem (see p. 5); it could be proved in a way entirely similar to the proof by Jacobi (see Krein [41]). However, it could be also obtained by means of Liénard–Chipart method, proceeding from the fact, that if $g = g_1 g_2$ then $g_\delta = g_{1\delta}g_2+g_1g_{2\delta}$.

It is easy to obtain the identity [41]

$$\mathcal{B}(g, ig_\delta; x_0,\dots,x_{n-1}) = \frac{1}{n}\mathcal{H}(g'; x_0,\dots,x_{n-2})+\frac{1}{n}|g'(x)|\,|\bar{g}'(x)|$$

from which the Cohn [12] theorem follows

XXII. *The number of roots of the symmetrical polynomial $g(z)$ inside the unit circle equals the number of different roots of its derivative $g'(z)$ outside this circle.*

Cohn has obtained this theorem with the help of the Rouché [62] theorem and many tedious "continuity considerations".

From the identity

$$\mathcal{H}(f+\lambda f^*; x_0,\dots,x_{n-1}) = (1-\lambda\bar{\lambda})\mathcal{H}(f; x_0,\dots,x_{n-1})$$

it follows that, when $|\lambda| < 1$, the polynomials $f(z)$ and $f(z)+\lambda f^*(z)$ have the same number of the roots inside the unit circle.† Wherefrom, with

$$\lambda = -\frac{a_n}{a_0}$$

the algorithmic rule by Schur [65] is obtained:

XXIII. *A polynomial* $f(z) = a_0 z^n + \cdots + a_n$ *has all its roots inside the unit circle, when, and only when,* $|a_n| < |a_0|$ *and the polynomial* $f_1(z)$, *determined from the equality*

$$z f_1(z) = \bar{a}_0 f(z) - a_n f^*(z)$$

possesses the same property.

Immediately from this rule the Eneström [14], [15] theorem follows, which is often named the Kakeya [40] theorem:

XXIV. *If* $a_0 > a_1 > \cdots > a_n > 0$, *then all roots of the equation*

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

are inside the unit circle.

This theorem is well known in its applications to the theory of functions (see Landau [46a], p. 20; [46b], p. 26).

Cohn [12] expanded the Schur rule, giving algorithms by means of which the numbers of roots of a polynomial outside and inside the unit circle can be calculated.

§ 4. ASSOCIATION WITH SOME PROBLEMS OF FUNCTION THEORY

Let us show that the problems touched upon in our paper have a close relation to that part of the theory of the functions of a complex variable which is known as the theory of bounded functions.

† Cohn [12] obtains the same proposition as a direct corollary of the Rouché theorem, for, on the unit circle, $|f(z)| = |f^*(z)|$; the proof given in the text has the advantage of being purely algebraic.

If $f(x) = h(x) - wg(x)$, where g and h are real polynomials and w is a complex number, then

$$H(f) = \frac{w - \bar{w}}{i} B(h, g).$$

Let the roc
definite, an
then the for
upper or i
Theorem X
$\frac{h(x)}{g(x)} = w$ li
words, the
which n-pl
possible to
$h(x)$ and $g($
Let us sh
related by
Bezoutiant
where z_1, z
half-plane.
$\sum_{k,j=1}^{n} \frac{h_k \bar{g}_j -}{z_k -}$
From here
are positiv
XXV. *In c*
$w_1, \ldots, w_n$ *i*
necessary
be positive
Moreov
functions

Let the roots of the polynomials h and g alternate; then the form $B(h,g)$ is definite, and, not violating generality, it could be thought to be positive. But then the form $H(f)$ would be positive or negative depending on w being in the upper or in the lower half-plane. Consequently, because of the Hermite Theorem XV, all roots of the equation $f(x) = 0$, that is also of the equation $\frac{h(x)}{g(x)} = w$ lie together with w in the upper or the lower half-plane. In other words, the function $w = \frac{h(x)}{g(x)}$ is a Nevanlinna–Pick function, that is, a function which n-ply maps the upper half-plane again into the upper half-plane. It is possible to show that this condition is also sufficient to ensure that the roots of $h(x)$ and $g(x)$ alternate (see Pick [61]).

Let us show that the values of the function w in different complex points are related by remarkable inequalities. Let us make a transformation in the Bezoutiant $B(h,g;x_0,\ldots,x_{n-1})$

$$x_k = \xi_1 z_1^k + \xi_2 z_2^k + \cdots + \xi_n z_n^k \qquad (k = 0,1,\ldots,n-1)$$

where $z_1, z_2, \ldots, z_n$ are arbitrary but distinct complex numbers from the upper half-plane. Then, the Bezoutiant $B(h,g)$ becomes

$$\sum_{k,j=1}^{n} \frac{h_k\bar{g}_j - \bar{h}_j g_k}{z_k - \bar{z}_j}\xi_k\bar{\xi}_j = \sum_{k,j=1}^{n} \frac{w_k - \bar{w}_j}{z_k - \bar{z}_j}(\xi_k g_k)(\overline{\xi_j g_j})$$

$$(h_k = h(z_k),\ g_k = g(z_k),\ w_k = w(z_k)).$$

From here it follows that all determinants

$$\left|\frac{w_k - \bar{w}_j}{z_k - \bar{z}_j}\right|_1^p \qquad (p = 1,\ldots,n)$$

are positive. Pick [61] has proved the following theorem:

XXV. *In order that there exist a rational function w of degree n taking values $w_1,\ldots,w_n$ in the points $z_1,\ldots,z_n$, and mapping the upper half-plane into itself, it is necessary and sufficient that all the determinants*

$$\left|\frac{w_k - \bar{w}_j}{z_k - \bar{z}_j}\right|_1^p \qquad (p = 1,\ldots,n)$$

be positive.

Moreover, he has proved that this system is preserved if the class of rational functions discussed is enlarged to the class of functions which are regular in the upper half-plane and which are not reducible to rational functions of degree lower than n.†

Similar results can be obtained by considering the symmetrical polynomials $g(z)$ and $h(z)$, with alternating roots, and, together with them, also the functions which map the unit circle into the upper half-plane.

As it is known, Carathéodory [8] first studied such functions (see also, Herglotz [27]).

Finally, the problems discussed are closely related to the Schur theory of functions which map the unit circle into a unit circle. Indeed, if a polynomial $f(z)$ has all roots inside the unit circle, i.e.

$$f(z) = a_0(z-\alpha_1)\cdots(z-\alpha_n) \qquad (|\alpha_k| < 1,\, k = 1,\dots,n),$$

then

$$f^*(z) = \bar{a}_0(1-\bar{\alpha}_1 z)\cdots(1-\bar{\alpha}_n z),$$

and, consequently, the function

$$F(z) = \frac{f(z)}{f^*(z)} = \frac{a_0}{\bar{a}_0}\prod_{k=1}^{n}\frac{z-\alpha_k}{1-\bar{\alpha}_k z} = c_0 + c_1 z + \cdots$$

maps the unit circle into the n-fold unit circle. This proposition can be reversed.

If, now, with the generating expression

$$\frac{1-F(x)\bar{F}(y)}{1-xy} = \sum_{k,l=0}^{\infty} C_{kl}x^k y^l$$

one constructs an infinite Hermitian form

$$\sum_{k,l=0}^{\infty} C_{kl}x_k\bar{x}_l$$

then it is easy to show that this form, by means of the transformation

† These results by Pick could be slightly generalized, namely, it is possible to establish: In order that there exist a rational function w of degree $\leqslant n$ or a general regular function in the upper half-plane, taking the values $w_1, w_2, \dots w_n$ in the points $z_1, z_2, \dots, z_n$ $(I(z_k) > 0)$ and mapping the upper half-plane into itself, it is necessary and sufficient that the Hermitian form

$$\sum_{k,j=1}^{n}\frac{w_k-\bar{w}_j}{z_k-\bar{z}_j}\xi_k\bar{\xi}_j$$

be non-negative. Observe, also, that Pick's investigations have been considerably deepened in a particular direction by the work of Nevanlinna [52], [53].

$$u_k = \ulcorner x^k f^*(x) \urcorner \qquad (k = 0, 1, \dots)$$

goes over into the form $\mathscr{H}(f; u_0, \ldots, u_{n-1})$ which is positive. Thus, the form

$$\sum_{k,l=0}^{\infty} C_{kl} x_k \bar{x}_l$$

is non-negative and has rank n. Schur [65] has proved a more general proposition:

In order that a function $F(z) = c_0 + c_1 z + \cdots$ be regular in the unit circle and map it into itself, it is necessary and sufficient that the form

$$L = \sum_{k,l=0}^{\infty} C_{kl} x_k \bar{x}_l$$

be non-negative. If the form L is non-negative and has finite rank n; then the function F(z) is always represented as $\frac{f(z)}{f^*(z)}$, where $f(z)$ is a polynomial of degree n, all roots of which lie inside the unit circle.

It is in connection with this theorem that Schur arrived at the result mentioned on page 34.

It is possible to pose a problem:

What characteristic function-theoretical properties has the quotient $\frac{f(z)}{f^*(z)}$ in the case in which $f(z)$ has a given number of roots inside the unit circle, with the remaining roots outside of this circle? Here T. Takagi [76], [77] has obtained an interesting result and, independently of him, N. Achiezer [1].

Finally, let us observe, concluding this paper, that many propositions with respect to the roots of polynomials, as for example, Borchardt's theorem, Hurwitz' theorem, etc. could be generalized to the case of integral transcendental functions.

Grommer first obtained fundamental results in this respect.

Grommer's work was simplified and partially complemented by Kritikos, N. G. Tschebotarev, A. F. Kravchuk and Fujiwara (see the additional list of literature).

Note by N. I. Achiezer with respect to his paper. "On a 'minimum' problem of function theory and on the number of roots of an algebraic equation which are inside the unit circle".†

Thanks to the courtesy of M. G. Krein, who drew my attention to the papers by Takagi "On an Algebraic Problem Related to an Analytic Theorem of Carathéodory and Féjèr and on an Allied Theorem of Landau"‡ and "Remarks on an Algebraic Problem", § I have learned that the problem which I discussed in §§ 2–3 of the paper named in the title of this note has been studied before me in the above mentioned works by Takagi.

The method which I have used differs from Takagi's method and, as it seems to me, is of some interest, for I proceed exclusively from the principle of the argument and discuss not only the problem with the Carathéodory–Féjèr

† *Communications of the Academy of Science, U.S.S.R.* (1931), No. 9, pp. 1169–1189.
‡ *Japanese Journal of Mathematics*, Vol. I (1924), pp. 83–93.
§ Ibid., Vol. II (1925), pp. 13–17.

conditions, but, also, with the Pick (§5) conditions, and I also take some conditions from the theorem from which Takagi proceeds.†

However, I, as also Takagi in his first paper, missed the possibility of an exceptional case, which Takagi calls irregular in his second paper, in which the rational fractions, discussed in my paper, do not exist.

For an irregular case, the study which Takagi gave in his second paper, the lemma and theorem of my § 2 and the similar statements of my § 5 are no longer valid, as well as the propositions corresponding to these statements in the first paper by Takagi.

Nevertheless, the basic Theorem 3, the reasoning of § 4, and Theorem 6 of my paper are valid also for the irregular case; this could be easily proved by means of considerations of continuity, as Takagi also points out.

N. ACHIEZER

† In his study Takagi assumes as known certain results by Carathéodory–Féjèr, as well as the algebraic theorems by Schur–Cohn, which I obtain using all the time the same method.

Index of the literature

[1] N. Achiezer, On a minimum problem in the theory of functions and on the number of roots of an algebraic equation which lie within the unit circle, *Izv. Akad. Nauk S.S.S.R.* (1931), 1169–1189 (37) (in Russian).

[2] I. Baur, Ueber die verschiedenen Wurzeln einer algebraischen Gleichung, *Math. Ann.* **50** (1898), 241–246 (6).

[3] Biehler, Sur une classe d'équations algébriques dont toutes les racines sont réelles, *J. Reine Angew. Math.* **87** (1879), 350–352 (27).

[4] C. Borchardt, Développements sur l'équation à l'aide de laquelle on détermine les inégalites séculaires du mouvement des planètes, *J. Math. Pures Appl.* **12** (1847), 50–67 (5).

[5] C. Borchardt, Bemerkung ueber die beiden vorstehenden Aufsaetze, *J. Reine Angew. Math.* **53** (1857), 281–283 (5).

[6] C. Borchardt, Remarque relative à la note precedente, *J. Reine Angew. Math.* **53** (1857), 367–368.

[7] F. Brioschi, Sur les séries qui donnent le nombre de racines réelles des équations algèbriques à une ou à plusieurs inconnues. *Nouv. Ann. de Math.* **15**: ser. 1 (1856), 264–286 (10, 12).

[8] C. Carathéodory, Ueber den Variabilitaetsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen, *Rend. Circ. Mat. Palermo* **32**: 2 (1911), 193–217 (36).

[9] A. Cauchy, Calcul des indices des fonctions, *J. École Polytechn.* **15**: Cah. 25 (1737), 176–229 (17, 23).

[10] A. Cayley, Nouvelles Recherches sur les fonctions de M. Sturm, *J. Math. Pures Appl.* **13** (24).

[11] A. Cayley, Note sur la méthode d'élimination de Bezout, *J. Reine Angew. Math.* **53** (1857), 366–376 (13).

[12] A. Cohn, Ueber die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise, *Math. Z.* **14** (1922), 110–148 (30, 30, 34, 34, 34).

[13] G. Darboux, Mémoire sur le théorème de Sturm, *Bull. Sci. Math.* **8** (1875), 56–63, 92–112 (17, 24).

[14] G. Eneström, Harledning af en allmän formel för antalet pensionärer som vid en godtycklig tidpunkt förifinn's inom en sluten pensions, *Kass. Öfversigt af Kongl. Vetenskaps-Akademiens Förhandingar* **50** (1893), 405–415 (34).

[15] G. Eneström, Remarque sur un théorème relatif aux racines de l'équation $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$ ou tous les coefficients a_i sont réelles et positifs, *Tôhoku Math. J.* **18** (1920), 34–36 (34).

[16] G. Frobenius, Ueber das Traegheitsgesetz der quadratischen Formen, *S-B. preuss. Akad. Wiss.* (1894), 241–256 Maerz und 407–431 Mai (6, 13, 21, 22, 24).

[17] M. Fujiwara, Ueber die Wurzeln der algebraischen Gleichungen, *Tôhoku Math. J.* **8** (1915), 78–85 (5, 29).

[18] M. Fujiwara, Ueber die algebraischen Gleichungen, deren Wurzeln in einem Kreise oder in einer Halbebene liegen, *Math. Z.* **24** (1926), 161–169 (25, 29, 30, 30).

[19] M. Fujiwara, Ueber die Bezoutiante zweier Polynome. *Japan. J. Math.* **3** (1925), 9–12.

[19a] M. Fujiwara, Ueber die Nullstellen der ganzen Funktionen vom Geschlecht Null und Eins, *Tôhoku Math. J.* **25** (1925), (29).

[20] D. Grave, On Certain Properties of Hesse's Covariant [Hesse, Univ. Izv. Kiev, 1903, No. 6, pp. 1–9] (33) (in Russian).

— 40 —

[21] D. Grave, Elements of Higher Algebra, 1st ed., Kiev (1914), 691 pp. (23, 27, 29) (in Russian).

[22] S. Gundelfinger, Ueber die Transformationen einer quadratischen Form in eine Summe von Quadraten, *J. Reine Angew. Math.* **91** (1881), 221–237 (6).

[23] E. Gursa, Course of mathematical analysis, Vol. II, part I, 2nd ed., 1933, 99 pp. (30) (in Russian).

[24] K. Hattendorff, Bemerkungen zu dem Sturm'schen Satze, *Nachr. Ges. Wiss.*, Goettingen (1873), 779–784 (10).

[25] K. Hattendorff, *Die Sturm'schen Funktionen*, 2 Aufl., Hannover (1874), 68 pp. (10, 21).

[26] O. Haupt, *Einfuehrung in die Algebra*, B.B.I, II, 1 Aufl., Leipzig (1929), 663 pp. (17).

[27] G. Herglotz, Ueber Potenzreihen mit positivem reellen Teil im Einheitskreis. *Ber. Verh. saechs. Akad., Leipzig*, **8** (1911), 501–509 (36).

[28] G. Herglotz, Ueber die Wurzelanzahl algebraischer Gleichungen innerhalb und auf dem Einheitskreis, *Math. Z.* **19** (1923), 26–34 (31, 32, 33).

[29] C. Hermite, Sur l'extension du théorème de M. Sturm à un systeme d'équations simultanées. *C.R. Acad. Sci. Paris*, **35** (1852), 52–54 (6, 11, 12, 12).

[30] C. Hermite, Remarques sur le théorème de M. Sturm, *C. R. Acad. Sci., Paris*, **36** (1853), 294–297 (6, 8, 10).

[31] C. Hermite, Extrait d'une lettre de Mr. Ch. Hermite de Paris à Mr. Borchardt de Berlin, sur le nombre des racines d'une équation algébrique comprises entre des limits données, *J. Reine Angew. Math.* **52** (1856), 39–51 (6, 7, 17, 20, 24, 27, 27, 32).

[32] C. Hermite, Sur l'indice des fractions rationnelles, *Bull. Soc. Math. France* **7** (1879), 128–131 (27).

[33] A. Hurwitz, Ueber die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt, *Math. Ann.* **46** (1895), 273–284; in French: *Nouv. Ann. de Math.* **15**: Sér. 3 (1896), 108–126 (15, 18, 19, 28).

[34] G. Jacobi, Theoremata nova algebraica circa systema duarum aequationum inter duas variabliles propositarum. *J. Reine Angew. Math.* 14 (1835), 281–288 (12).

[34a] G. Jacobi, De eliminatione variabilis e duabus aequationibus algebraicis. *J. Reine Angew. Math.* 15 (1836), 285 (24).

[35] G. Jacobi, Ueber eine elementare Transformation eines in Bezug auf jede von zwei variablen Systemen linearen und homogenen Ausdrucks. *J. Reine Angew. Math.* 53 (1857), 265–270 (6).

[36] G. Jacobi, Ueber einen algebraischen Fundamentalsatz und seine Anwendungen. *J. Reine Angew. Math.* 53 (1857), 275–280 (7).

[37] J. Jamamoto, Eine Bemerkung ueber algebraische Gleichungen. *Tôhoku Math. J.* 33: 1, 2 (1930), (20).

[38] R. Jentzsch, Ueber eine Klasse von algebraischen Gleichungen mit lauter reellen Wurzeln. *Arch. d. Math. u. Phys.* 17 (1910), 105–107 (27).

[39] F. Joachimstal, Bemerkungen ueber den Sturm'schen Satz, *J. Reine Angew. Math.* 48 (1854), 386–416 (8, 21).

[40] S. Kakeya, On the Limits of the Roots of an Algebraic Equation with Positive Coefficients. *Tôhoku Math. J.* 2 (1912), 140–142 (34).

[41] M. Krein, On the Theory of Symmetric Polynomials. *Rec. Math., Moscow*, 40 (1933), 271–283 (29, 32, 33, 34) (in Russian).

[41a] M. Krein and M. Naimark, Ueber eine Transformation der Bezoutiante, die zum Sturm'schen Satze fuehrt. *Comm. de la Soc. Math. de Kharkov*, ser. 4, t. 10 (1934), 33–39 (23).

[42] L. Kronecker, Sur le théorème de Sturm, *C. R. Acad. Sci. Paris* 68 (1869), 1078–1082 (10).

[43] L. Kronecker, Ueber die verschiedenen Sturm'schen Reihen und ihre gegenseitigen Beziehungen, *Monatsber. preuss. Akad. Wiss.* a) Februar (1873), 117–154, b) Februar (1878), 95–121 (7, 10, 10, 21, 23, 24).

[44] Lagrange, Recherches sur la méthode de maximis et minimis. *Oeuvres*, t. 1, Paris (1867) 3–20 (24).

[45] E. Laguerre, Sur quelques théorèmes de M. Hermite, extrait d'une lettre adressée a M. Borchardt. *J. Reine Angew. Math.* 89 (1880), 339–342 (27).

[46] E. Landau, Darstellung und Begruendung einiger neuerer Ergebnisse der Funktionentheorie. a) 1 Aufl. Berlin, 1916, 110 pp., and b) 2 Aufl. Berlin, 1929, 122 pp. (34).

—41—

[47] A. Liénard, Condition pour qu'une équation algébrique ait toutes ses racines négatives ou à partie réelle négative. *Revue de math. spéc.* 21 (1911), 153–155, 177–180 (29).

[48] A. Liénard et M. Chipart, Sur le signe de la partie réelle des racines d'une équation algébrique. *J. Math. Pures Appl.* 10: sér. 6 (1914), 291–346 (25, 26, 29).

[49] A. Markov, On Functions Obtained by Transformations of Series into Continued Fractions. *Trans. Acad. Sci.*, Supplement to 74, (1894), 1–30 (21) (in Russian).

[50] A. Markov, On a Theorem in Higher Algebra. *Trans. St. Petersburg Math. Soc.* (1899), 26 (33) (in Russian).

[51] A. Markov, Some Generalizations of the Same Theorem. Ibid. (1899), 36–40 (33) (in Russian).

[52] R. Nevanlinna, Asymptotische Entwicklungen beschraenkter Funktionen und das Stieltjes'sche Momentenproblem, *Ann. Acad. Sc. Fenn.* 18: 5 (1922), 1–53, Ser. A (36).

[53] R. Nevanlinna, Ueber beschraenkte analytische Funktionen. Ibid. 32: 7 (1929), 1–75, Ser. A (36).

[54] L. Orlando, Sul problema di Hurwitz, *Atti. Accad. Naz. Lincei* 19 (1910), 801–805 (29).

[55] L. Orlando, Nuovo osservazioni sul problema di Hurwitz. Ibid. 19 (1910) (29).

[56] L. Orlando, Sull'equazione alle semisomme e sul teorema di Hurwitz. Ibid. 19: 2 (1910), 317–321 (29).

[57] L. Orlando, Sopra alcune question relative al problema di Hurwitz. Ibid. 19: 2 (1910), 430–434 (29).

[58] L. Orlando, Sulla dimonstrazione elementare del teorema di Hurwitz. Ibid. 20: 1 (1911), 742–745 (29).

[59] L. Orlando, Sul problema di Hurwitz relativo alle parti reali delle radici di un'equazione algebrica. *Math. Ann.* **71** (1911), 233–245; here the author presents in detail his earlier investigations of this problem (29).

[60] K. Petr. Die symmetrischen Zahlensysteme und der Satz von Sturm, *Bull. intern.* **11** (1906), 14–34; translation from the Czech into German, original in *Rozpravy česke Akademie cisarě Frantiska Josefa pro vědy, slovenost a umĕnl Prag.* **15**: 2 (1906), 1–19 (22).

[61] G. Pick, Ueber die Beschraenkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden, *Math. Ann.* **77** (1916), 7–23 (35, 35).

[62] E. Rouché, Mémoire sur la série de Lagrange, *J. Ecole Polytechn.* **22**: Cah. 39 (1862), 217–218 (30, 34).

[63] E. Routh, *A Treatise on the Stability of a Given State of Motion.* London (1877), 108 pp. (29).

[64] E. Routh, *The Advanced Part of a Treatise on the Dynamics of a System of Rigid Bodies.* II Part, 5 Ed., London (1892), 431 pp. (29).

[65] I. Schur, Ueber Potenzreihen, die im Innern des Einheitskreises beschraenkt sind, *J. Reine Angew. Math.* **148** (1918), 122–145 (30, 30, 34, 37).

[66] I. Schur, Ueber algebraische Gleichungen, die nur Wurzeln mit negativen Realteilen besitzen, *Z. Angew. Math. Mech.* **1** (1921), 307–311 (29).

[67] A. Stodola, Sur le réglement des turbines. *Schweizer Bau-Zeitung* **23**: 17, 18 (29).

[68] C. Sturm, Mémoire sur la résolution des équations numériques, *Bull. de Ferussac*, **11** (1829), or Mémoires présentés par divers savants à l'Acad. sci. France 6 (1835), 273–318; a German translation with notes by A. Loewy is available in Ostwalds "Klassiker der exakten Wiss.", Nr. 143, Leipzig (1904), 66 pp. (5, 21).

[69] C. Sturm et J. Liouville, D'un Théorème de Mr. Cauchy relatif aux racines imaginaires des Equations. *J. Math. Pures Appl.* **1** (1836), 278–289 (27).

[70] C. Sturm, Autres démonstrations du même Théorème. *J. Math. Pures Appl.* **1** (1836), 290–308 (27).

[71] C. Sturm, Démonstration d'un Théorème d'algèbre de Mr. Sylvester, *J. Math. Pures Appl.* **7** (1842), 356–368 (10, 23).

[72] J. Sylvester, Memoir on Rational Derivation from Equations of Co-existence, that is to say, a new and Extended Theory of Elimination, *Philos. Mag.* **15** (1839), 429–435 (7, 10).

[73] J. Sylvester, A Demonstration of the Theorem that every Homogeneous

Quadratic Polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares, *Philos. Mag.* 4 (1852), 138–142 (6).

[74] J. Sylvester, On a Theory of the Syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's Functions, and that of the greatest Algebraical Common Measure, *Philos. Trans. Roy. Soc. London* **143** (1853), 407–548 (7, 8, 10, 13, 17, 19).

[75] J. Sylvester, On a remarkable Modification of Sturm's Theorem, *Philos. Mag.* (1853), 446 (23).

[76] T. Takagi, On an Algebraic Problem related to an Analytic Theorem of Carathéodory and Féjèr and on an Allied Theorem of Landau, *Japan. J. Math.* **1** (1924), 83–93 (37).

[77] T. Takagi, Remarks on an Algebraic Problem, *Japan. J. Math.* **2** (1925), 13–17 (37).

Supplementary list of the literature relevant to §4

M. Fujiwara, Ueber die Nullstellen der ganzen Funktionen vom Geschlecht Null und Eins. *Tôhoku Math. J.* **25** (1925), 27–35.

J. Grommer, Ganze transzendenten Funktionen mit lauter reellen Nullstellen. *J. Reine Angew. Math.*, Bd. 144 (1914).

M. Krawchuk, Algebraic Studies of Analytic Functions. *Transact. of the Physics and Mathematics Department*, Vol. XII, 1st issue (in Ukrainian).

N. Kritikos, Ueber ganze transzendente Funktionen, usw., *Math. Ann.*, Bd. 81.
N. Tschebotareff, Ueber die Realitaet von Nullstellen ganzer transzendenten Funktionen, *Math. Ann.* Bd. 99 (1928), Ss. 660–686.

List of text-books

1. Bieberbach-Bauer, *Vorlesungen ueber Algebra*, Leipzig (1928), 334 pp.
2. R. Fricke, *Lehrbuch der Algebra*, Braunschweig (1924), 468 pp.
3. D. Grave, *Elements of Higher Algebra*, Kiev (1914), 691 pp. (in Russian).
4. O. Haupt, *Einfuehrung in die Algebra*, B. B. I, II, Leipzig (1929), 663 pp.
5. E. Netto, *Vorlesungen ueber Algebra*, B. I. Leipzig (1896), 388 pp.
6. E. Pascal, *Repertorium der hoeheren Analysis*, B. 1, 2, Aufl. Leipzig und Berlin (1910), 527 pp.
7. J. Serret, *Cours d'algèbre supérieure*, t. I. 3 éd., Paris (1866), 613 pp.

End of Translation

Note by M. G. Krein: With respect to the Schur and Takagi problems and their further development see the paper Analytic Properties of Schmidt Pairs for a Hankel Operator and the Generalized Schur–Takagi Problem (Russian) by V. M. Adamjan, D. Z. Arov and M. G. Krien, *Mat. Sb* 86 (128), No. 1 (1971), 39–75. (AMS Translation: *Math. USSR Sbornik*, Vol. 15 (1971), No. 1, 32–73.)

Acknowledgement

The translators wish to express their sincere thanks to the Lenin Library, Moscow, for a microfilm copy of the original text, and to Professor M. G. Krein for his kind permission to publish this translation. They are also grateful to Professor Krein for a list of corrections to the original text, which have been incorporated into the translation, and for the above note on the Schur–Takagi problem.

Comment by the Editors: The well-known Matrix Theory book by F. R. Gantmacher contains a reference to a 1936 article by M. G. Krein and M. A. Naimark, but with little publication information; the article seems not to have been published before in any form making it accessible outside the Soviet Union. We are greatly indebted to J. L. Howland for, first, securing a Russian language copy of this article, second, securing its translation into English, and third, securing permission from Professor Krein for its publication (including some corrections provided by Professor Krein). *Linear and Multilinear Algebra* is honored to be able to publish this unique article.

Dr Howland's translation preserves the paging and footnoting of the Russian text, displaying the original page numbers as they occur and using a horizontal line to separate the main text from the footnotes on each original page.