ON THE ROOTS OF THE CHARACTERISTIC EQUATION OF A LINEAR SUBSTITUTION

BY

T. J. I'A BROMWICH
in GALWAY, Ireland.

1. The equation in λ

$$\begin{vmatrix} a_{1,1}-\lambda, & a_{1,2}, & a_{1,3}, & \dots, & a_{1,n} \\ a_{2,1}, & a_{2,2}-\lambda, & a_{2,3}, & \dots, & a_{2,n} \\ a_{3,1}, & a_{3,2}, & a_{3,3}-\lambda, & \dots, & a_{3,n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n,1}, & a_{n,2}, & a_{n,3}, & \dots, & a_{n,n}-\lambda \end{vmatrix} = 0$$

has been discussed by many writers; the following results are well known.

The roots are real in case all the numbers a are *real* and such that $a_{r,s} = a_{s,r}$; that is, if the matrix of a's is symmetric.[1]

The roots have the absolute value unity, if the matrix of a's belongs to a real orthogonal substitution.[2]

The roots are pure imaginaries or zero, in case the a's are *real* and $a_{r,r} = 0$, $a_{r,s} = -a_{s,r}$; that is, if the matrix of a's is alternate.[3]

However, in spite of these results relating to special types of the matrix a, nothing was known of the nature of the roots for a general

[1] CAUCHY, 1829.

[2] BRIOSCHI, 1854.

[3] WEIERSTRASS, 1879.

matrix, until the problem was attacked by BENDIXSON[1] in 1900; he obtained upper and lower limits for the magnitude of the real and imaginary parts of the roots, taking all the numbers a to be real. The extension to the case of complex numbers a was made by HIRSCH[2] in 1902.

In what follows, we shall obtain narrower limits for the imaginary parts of the roots; incidentally, we also obtain BENDIXSON's and HIRSCH's limits for the real parts of the roots.

2. Take, in the first instance, all the a's to be *real;* and then write

$$\left.\begin{array}{lll} b_{r,r} = a_{r,r}, & b_{r,s} = \quad b_{s,r} = \frac{1}{2}(a_{r,s} + a_{s,r}), & \\ c_{r,r} = 0, & c_{r,s} = -c_{s,r} = \frac{1}{2}(a_{r,s} - a_{s,r}), & \\ A = \Sigma a_{r,s} x_r y_s, & B = \Sigma b_{r,s} x_r y_s, & C = \Sigma c_{r,s} x_r y_s. \end{array}\right\} \scriptstyle (r,s=1,2,\ldots,n)$$

It is now obvious that $A = B + C$, and that the bilinear forms B, C, are, respectively, symmetric and alternate. Following FROBENIUS, let us also write E for the unit form $\Sigma x_r y_r$ and let $|A - \lambda E|$ denote the determinant written out at the beginning of § 1, while $|B - \lambda E|$, $|C - \lambda E|$, stand for similar determinants with b's, c's in place of a's.

Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$, are the (real) roots of $|B - \lambda E|$, it is then known from a theorem due to WEIERSTRASS[3] that a *real* linear substitution can be found which, when applied to the x's and y's, reduces B to the form $B_1 = \Sigma \lambda_r x_r y_r$, while it leaves E unchanged. This substitution will change C into C_1, another alternate form with real coefficients; but it will not alter the roots of the fundamental equation. Thus the equation $|B_1 + C_1 - \lambda E| = 0$ has the same roots as $|A - \lambda E| = 0$.

Suppose now that $\lambda = \alpha + i\beta$ is one of these roots; then the bilinear form $B_1 + C_1 - (\alpha + i\beta)E$ has the rank[4] $(n - 1)$ at most. Consequently values of $x_1, x_2, \ldots, x_n$ can be chosen which make the form zero, whatever

---

[1] Öfversigt af K. Vet. Akad. Förh. Stockholm, 1900, Bd. 57, p. 1099; Acta Mathematica, t. 25, 1902, p. 359.

[2] Acta Mathematica, l. c., p. 367.

[3] Berliner Monatsberichte, 1858; Ges. Werke, Bd. 1, p. 243.

[4] Rang, according to FROBENIUS.

values may be taken for $y_1, y_2, \dots, y_n$; naturally, the values for the x's will usually be complex, and some of them must be complex, unless β is zero. Write for these special values

$$x_r = p_r + iq_r, \qquad (r=1,2,\dots,n)$$

and let us choose for the y's the conjugate complex numbers

$$y_r = p_r - iq_r, \qquad (r=1,2,\dots,n)$$

it being understood that p_r and q_r are real. With these values for x_r, y_r, we find

$$B_1 = \Sigma\lambda_r(p_r^2 + q_r^2), \qquad E = \Sigma(p_r^2 + q_r^2); \qquad (r=1,2,\dots,n)$$

further

$$x_r y_s - x_s y_r = -2i(p_r q_s - p_s q_r),$$

so that C_1 becomes a pure imaginary. But, according to what we have already explained,

$$\Sigma\lambda_r(p_r^2 + q_r^2) + C_1 - (\alpha + i\beta)\Sigma(p_r^2 + q_r^2) = 0;$$

thus, since C_1 is imaginary only, we have

$$\Sigma\lambda_r(p_r^2 + q_r^2) - \alpha\Sigma(p_r^2 + q_r^2) = 0.$$

Hence

$$\alpha = \frac{\Sigma\lambda_r(p_r^2 + q_r^2)}{\Sigma(p_r^2 + q_r^2)},$$

and consequently α *lies between the greatest and least of the numbers* λ_1, $\lambda_2, \dots, \lambda_n$, which is one of BENDIXSON's results (l. c. Theorem II).

We proceed next to obtain a corresponding theorem for β. Let us suppose that the non-zero roots of the equation $|C - \lambda E| = 0$ are given by $\lambda = \pm i\mu_1, \pm i\mu_2, \dots, \pm i\mu_\nu$, where $2\nu \leqq n$; so that there are $(n - 2\nu)$ zero roots of this equation. By a theorem of WEIERSTRASS,[1] stated in § 1, the numbers $\mu_1, \mu_2, \dots, \mu_\nu$ are all real; and they may be supposed positive without loss of generality. Further the invariant-factors of the determinant $|C - \lambda E| = 0$ are all *linear*.[1] It is then possible to find a

[1] WEIERSTRASS, Berliner Monatsberichte, 1870; Ges. Werke, Bd. 3, p. 139.

real linear substitution, which, when applied to the x's and y's, reduces C to the form

$$C_2 = \mu_1(x_1y_2 - x_2y_1) + \mu_2(x_3y_4 - x_4y_3) + \ldots + \mu_\nu(x_{2\nu-1}y_{2\nu} - x_{2\nu}y_{2\nu-1}),$$

but leaves E unchanged.[1] Owing to the nature of this substitution, B is changed to B_2 another bilinear form which is symmetric and has real coefficients. Then, just as in the last case, values of the x's can be chosen so that $B_2 + C_2 - (\alpha + i\beta)E = 0$, for all values of the y's. Let these values of the x's be given by

$$x_r = p_r + iq_r, \qquad (r=1, 2, \ldots, n)$$

and take

$$y_r = p_r - iq_r. \qquad (r=1, 2, \ldots, n)$$

Then

$$x_ry_r = p_r^2 + q_r^2, \qquad x_ry_s + x_sy_r = 2(p_rp_s + q_rq_s),$$

and consequently B_2 is real; but

$$x_ry_s - x_sy_r = -2i(p_rq_s - p_sq_r)$$

so that

$$C_2 = -2i[\mu_1(p_1q_2 - p_2q_1) + \ldots + \mu_\nu(p_{2\nu-1}q_{2\nu} - p_{2\nu}q_{2\nu-1})].$$

Hence, from the equation $B_2 + C_2 - (\alpha + i\beta)E = 0$ we deduce

$$\beta\Sigma(p_r^2 + q_r^2) = -2[\mu_1(p_1q_2 - p_2q_1) + \ldots + \mu_\nu(p_{2\nu-1}q_{2\nu} - p_{2\nu}q_{2\nu-1})].$$

But, in absolute value $2(p_1q_2 - p_2q_1)$ is not greater than $(p_1^2+q_2^2) + (p_2^2+q_1^2)$, and consequently

$$|\beta|\,\Sigma(p_r^2 + q_r^2) \leqq [\mu_1(p_1^2 + q_1^2 + p_2^2 + q_2^2) + \ldots + \mu_\nu(p_{2\nu-1}^2 + q_{2\nu-1}^2 + p_{2\nu}^2 + q_{2\nu}^2)].$$

From which it is clear that *the absolute value of β cannot exceed the greatest of the numbers $\mu_1, \mu_2, \ldots, \mu_\nu$;* which is obviously analogous to Bendixson's Theorem II. We shall now see that *this theorem usually gives narrower limits for β than Bendixson's Theorem I, and cannot give wider limits.*

---

[1] That such a reduction is possible is contained implicitly in Kronecker's work on the reduction of a single bilinear form. For an explicit treatment, see my papers, Proc. Lond. Math. Soc., vol. 32, 1900, p. 321, § 4; vol. 33, 1901, p. 197, § 3; American Journal of Mathematics, vol. 23, 1901, p. 235.

For, since $\pm i\mu_1, \pm i\mu_2, \ldots, \pm i\mu_\nu$ are the non-zero roots of $|\lambda E - C| = 0$, it follows that

$$\mu_1^2 + \mu_2^2 + \ldots + \mu_\nu^2$$

is equal to the coefficient of λ^{n-2} in the expanded form of the determinant; thus

$$\mu_1^2 + \mu_2^2 + \ldots + \mu_\nu^2 = \frac{1}{2} \Sigma c_{r,s}^2. \qquad (r, s = 1, 2, \ldots, n)$$

Hence, if g is the greatest coefficient in C, we have [1]

$$\mu_1^2 + \mu_2^2 + \ldots + \mu_\nu^2 \leqq \frac{1}{2} n(n-1) g^2.$$

Thus it will usually happen that the greatest of $\mu_1, \mu_2, \ldots, \mu_\nu$ is less than $g\left[\frac{1}{2} n(n-1)\right]^{\frac{1}{2}}$; and the greatest μ can never exceed this value, which is the limit given in BENDIXSON's Theorem I.

3. Suppose now that the numbers a are *complex:* and write a' to denote the complex number conjugate to a. Then write

$$\left.\begin{aligned} b_{r,s} &= \frac{1}{2}(a_{r,s} + a'_{s,r}), & b_{s,r} &= \frac{1}{2}(a_{s,r} + a'_{r,s}), \\ ic_{r,s} &= \frac{1}{2}(a_{r,s} - a'_{s,r}), & ic_{s,r} &= \frac{1}{2}(a_{s,r} - a'_{r,s}), \end{aligned}\right\} \qquad (r, s = 1, 2, \ldots, n)$$

so that,

$$b'_{r,s} = b_{s,r}, \qquad c'_{r,s} = c_{s,r}. \qquad (r, s = 1, 2, \ldots, n)$$

Further, put

$$A = \Sigma a_{r,s} x_r y_s, \qquad B = \Sigma b_{r,s} x_r y_s, \qquad C = \Sigma c_{r,s} x_r y_s, \quad (r, s = 1, 2, \ldots, n)$$

Then it is obvious that $A = B + iC$, and that the bilinear forms B, C are forms of HERMITE's type.

Suppose now that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the roots of $|B - \lambda E| = 0$; it is known that these roots are all real and that the invariant-factors of the determinant are linear.[2] It is then possible to find a linear substitution

---

[1] There are only $n(n-1)$ non-zero coefficients in C, because $c_{r,r} = 0$.

[2] CHRISTOFFEL, Crelle's Journal, Bd. 63, 1864, p. 252.

S (usually complex) such that when S is applied to the x's, and the conjugate substitution to the y's, the form B is reduced to $B_1 = \Sigma\lambda_r x_r y_r$, while E remains unchanged.[1] Further C is changed to C_1, another bilinear form of HERMITE's type, (in consequence of the relation between the substitutions on the x's and on the y's).

The determinantal equation then becomes $|B_1 + iC_1 - \lambda E| = 0$; thus, if a root is $\lambda = \alpha + i\beta$, we can choose the x's so as to make

$$B_1 + iC_1 - (\alpha + i\beta)E = 0,$$

whatever values we give to the y's. Suppose that these values for the x's are given by

$$x_r = p_r + iq_r, \qquad (r = 1, 2, \ldots, n)$$

and then take

$$y_r = p_r - iq_r = x'_r. \qquad (r = 1, 2, \ldots, n)$$

Thus

$$B_1 = \Sigma\lambda_r(p_r^2 + q_r^2), \qquad E = \Sigma(p_r^2 + q_r^2);$$

also, if $C_1 = \Sigma\gamma_{r,s} x_r y_s$, we have that $\gamma_{r,s} x_r y_s$ and $\gamma_{s,r} x_s y_r$ are conjugate complex numbers, because $\gamma_{s,r} = \gamma'_{r,s}$, $x_s = y'_s$, $y_r = x'_r$; further $\gamma_{r,r} x_r y_r$ is real; hence B_1, C_1, E are all three *real*. Consequently the relation

$$B_1 + iC_1 - (\alpha + i\beta)E = 0$$

gives $B_1 = \alpha E$, so that

$$\alpha = \frac{\Sigma\lambda_r(p_r^2 + q_r^2)}{\Sigma(p_r^2 + q_r^2)}.$$

Thus, just as in § 2, *α lies between the greatest and least of* $\lambda_1, \lambda_2, \ldots, \lambda_n$. This is HIRSCH's Theorem II.

But it is now clear that, if $\lambda = \mu_1, \mu_2, \ldots, \mu_n$ are the roots of $|C - \lambda E| = 0$, we can similarly transform C into the form $C_2 = \Sigma\mu_r x_r y'_r$, leaving E unchanged, while B becomes B_2 another HERMITE's form. Thus, by an exactly similar argument, we find that *β lies between the greatest and least of* $\mu_1, \mu_2, \ldots, \mu_n$; which is the extension to complex coefficients of the theorem proved in § 2 for real coefficients.

We proceed now to show the connection between these theorems and

---

[1] See for example § 6 of the first, or § 5 of the last, of my papers quoted above.

HIRSCH's Theorem I. Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the roots of the equation $|B - \lambda E| = 0$, by comparing coefficients of $\lambda^{n-1}, \lambda^{n-2}$, we find

$$\Sigma\lambda_r = \Sigma b_{r,r}, \qquad \Sigma\lambda_r\lambda_s = \Sigma(b_{r,r}b_{s,s} - b_{r,s}b_{s,r}). \qquad (r, s = 1, 2, \ldots, n)$$

Thus

$$\Sigma\lambda_r^2 = \Sigma b_{r,r}^2 + \Sigma b_{r,s}b_{s,r}.$$

Hence, if g_1 is the greatest absolute value of any coefficient in B, we have

$$\Sigma\lambda_r^2 \leqq ng_1^2 + n(n-1)g_1^2,$$

or

$$\Sigma\lambda_r^2 \leqq (ng_1)^2.$$

Now we have seen that α^2 is not greater than the greatest of $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$; and consequently α^2 is usually less than $(ng_1)^2$, while it can never be greater than this limit. That is, *α is not greater, numerically, than ng_1*. Similarly, if g_2 is the greatest absolute value of any coefficient in C, it can be proved that[1] *β is not greater, numerically, than ng_2*.

From the inequality proved above

$$\alpha^2 \leqq \Sigma b_{r,r}^2 + \Sigma b_{r,s}b_{s,r} \qquad (r, s = 1, 2, \ldots, n)$$

and the corresponding one

$$\beta^2 \leqq \Sigma c_{r,r}^2 + \Sigma c_{r,s}c_{s,r},$$

we find

$$\alpha^2 + \beta^2 \leqq \Sigma(b_{r,r}^2 + c_{r,r}^2) + \Sigma(b_{r,s}b_{s,r} + c_{r,s}c_{s,r}).$$

Now

$$b_{r,r}^2 + c_{r,r}^2 = a_{r,r}a'_{r,r},$$

and

$$b_{r,s}b_{s,r} + c_{r,s}c_{s,r} = \frac{1}{2}(a_{r,s}a'_{s,r} + a_{s,r}a'_{r,s}),$$

so that

$$\alpha^2 + \beta^2 \leqq \Sigma a_{r,r}a'_{r,r} + \Sigma a_{r,s}a'_{s,r}. \qquad (r, s = 1, 2, \ldots, n)$$

Thus, if g_3 is the greatest absolute value of any coefficient in A, we have

$$\alpha^2 + \beta^2 \leqq ng_3^2 + n(n-1)g_3^2,$$

[1] If it happens that the coefficients in C are pure imaginaries, so that $c_{r,r} = 0$, $c_{r,s} = -c_{s,r}$, it can be proved (as in § 2) that

$$|\beta| \leqq g_2\left[\frac{1}{2}n(n-1)\right]^{\frac{1}{2}}.$$

or

$$|\alpha + i\beta| \leqq ng_3.$$

That is, *the absolute value of* $(\alpha + i\beta)$ *is not greater than* ng_3.

The results

$$|\alpha| \leqq ng_1, \qquad |\beta| \leqq ng_2, \qquad |\alpha + i\beta| \leqq ng_3$$

constitute Hirsch's Theorem I, which is therefore included in the general theorem obtained previously.

4. I have also attempted to obtain some relation between the indices of the invariant-factors of $|A - \lambda E|$, and those of[1] $|\lambda B + \mu C|$; but hitherto I have not succeeded in finding any general theorem in this connection. The two following examples show that the relation (if there is one) is not very obvious.

If

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda, & 2, & 4 \\ 0, & 1-\lambda, & 6 \\ 0, & 0, & 1-\lambda \end{vmatrix}, \qquad \text{[\textit{One} invariant-factor } (\lambda - 1)^3]$$

then

$$|\lambda B + \mu C| = \begin{vmatrix} \lambda, & \lambda+\mu, & 2(\lambda+\mu) \\ \lambda-\mu, & \lambda, & 3(\lambda+\mu) \\ 2(\lambda-\mu), & 3(\lambda-\mu), & \lambda \end{vmatrix}. \qquad \text{[\textit{Three} invariant-factors } \lambda(\lambda^2 - 2\mu^2)]$$

Again if $|A - \lambda E| = \begin{vmatrix} a-\lambda, & -1 \\ 1, & -\lambda \end{vmatrix}$, then $|\lambda B + \mu C| = \begin{vmatrix} a\lambda, & -\mu \\ \mu, & 0 \end{vmatrix}$.

In this case both determinants have a squared invariant-factor if $a^2 = 4$; but if a has any other value, the first has two different invariant-factors $(\lambda^2 - a\lambda + 1)$, while the second has always a squared invariant-factor (μ^2).

Dublin, 11th October, 1904.

---

[1] It is obviously hopeless to use the invariant-factors of $|B - \lambda E|$ and $|C - \lambda E|$, because these are always *linear;* while $|A - \lambda E|$ may have invariant-factors of any degree up to n. In this paragraph the a's are supposed real, so that B and C are deduced from A according to § 2 (not § 3).