

ON THE LOCATION OF THE ROOTS
OF POLYNOMIALS*

BY J. L. WALSH

It is the object of this note to prove a number of results (particularly Theorems I, II, V, VI, below) concerning the roots of polynomials, generalizations of former results established by the writer. Our main new result is the following theorem.

THEOREM I. *Let (the interiors and boundaries of) the circles C_1, C_2, \dots, C_k whose centers are the points $\alpha_1, \alpha_2, \dots, \alpha_k$ be the respective loci of n_1, n_2, \dots, n_k roots of a variable polynomial $f(z)$ which has no other roots, where the circles C_i are all equal and their centers α_i all lie on a line parallel to the axis of reals. If the polynomial*

$$(1) \quad a_0 z^n + n a_1 z^{n-1} + n(n-1) a_2 z^{n-2} \\ + \dots + n(n-1) \dots 2 \cdot 1 a_n, \\ (n = n_1 + n_2 + \dots + n_k),$$

has only real roots, and if the circles C_i are sufficiently small, then the locus of the roots of the polynomial

$$F(z) = a_0 f(z) + a_1 f'(z) + a_2 f''(z) + \dots + a_n f^{(n)}(z)$$

consists of the circles C'_j which are equal to the circles C_i and whose centers are the roots of $F(z)$ when the roots of $f(z)$ are the points $\alpha_1, \alpha_2, \dots, \alpha_k$ of multiplicities n_1, n_2, \dots, n_k respectively.† Any circle C'_j which has no point in common with any of the other circles C'_m contains a number of roots of $F(z)$ equal to the multiplicity of its center as a root of $F(z)$ when the roots of $f(z)$ are the points α_i .

We shall later make clear the exact meaning of the words, "if the circles C_i are sufficiently small". Theorem I is to be proved by iteration of the following theorem.

* Presented to the Society, December 27, 1922.

† When the roots of $f(z)$ are the points $\alpha_1, \alpha_2, \dots, \alpha_k$, the roots of $F(z)$ are all collinear with the roots of $f(z)$. This is a well known theorem due to Hermite, which can easily be established by proving the result in succession for the polynomials $F_1(z), F_2(z), \dots, F_n(z)$ used below.

THEOREM II. *Let the circles C_1, C_2, \dots, C_k whose centers are the points $\alpha_1, \alpha_2, \dots, \alpha_k$ be the loci respectively of n_1, n_2, \dots, n_k roots of a variable polynomial $f(z)$ which has no other roots, where the circles C_i are all equal and their centers α_i all lie on a line parallel to the axis of reals. If c_1 is real, then the locus of the roots of the polynomial*

$$(2) \quad F_1(z) = f(z) + c_1 f'(z)$$

consists of the circles C'_j which are equal to the circles C_i and whose centers are the roots of $F_1(z)$ when the roots of $f(z)$ are the points $\alpha_1, \alpha_2, \dots, \alpha_k$ of multiplicities n_1, n_2, \dots, n_k respectively. Any circle C'_j which has no point in common with any other of the circles C'_m contains a number of the roots of $F_1(z)$ equal to the multiplicity of its center as a root of $F_1(z)$ when the roots of $f(z)$ are the points α_i .

Theorem I has already been established for the case $a_0 = a_1 = \dots = a_{e-1} = a_{e+1} = a_{e+2} = \dots = a_n = 0, a_e \neq 0$.*

Theorem I has also been proved with no restriction on (1) nor on the size of the circles C_i for the case $k = 1$.† The limiting case for $c_1 = \infty$ of Theorem II is also true, and is proved as a limiting case of I, Theorem VIII. For the case that the circles C_i of Theorem II are sufficiently small, Theorem II is contained in Theorem I.

The limiting case for $c_1 = \infty$ of Theorem II gives

* TRANSACTIONS OF THIS SOCIETY, vol. 24 (1922), pp. 31-69, Theorem X, p. 53. We shall refer to this paper as I.

† TRANSACTIONS OF THIS SOCIETY, vol. 24 (1922), pp. 163-180, Theorem VI. We shall refer to this paper as II.

The note in the COMPTES RENDUS of which II is the development was published and II itself was written before the publication of a paper by S. Kakeya, PROCEEDINGS OF THE PHYSICO-MATHEMATICAL SOCIETY OF JAPAN, (3), vol. 3 (1921), pp. 94-100. Kakeya treats the main theorem of II, and by essentially the methods of II, although he makes the restriction that the circular region C involved shall be bounded by a circle whose center is the origin. See also T. Takagi, PROCEEDINGS OF THE PHYSICO-MATHEMATICAL SOCIETY OF JAPAN, (3), vol. 3 (1921), pp. 175-179, who gives a proof of Theorem V of the present paper.

a theorem which is essentially a well known theorem due to Lucas:

If a circle contains all the roots of a polynomial, that circle contains all the roots of its derivative.

Theorem I contains the more general theorem of Lucas which deals with any derivative of the original polynomial, for it turns out that in this case the circles C_i are "sufficiently small" in the sense which we shall give to those words.

Let us proceed to the proof of Theorem II. In the case $c_1 = 0$, the hypothesis and conclusion are the same; we turn to the case $c_1 \neq 0$. The roots of $F_1(z)$ are the roots of

$$(3) \quad -\frac{1}{c_1} = \frac{f'(z)}{f(z)} = \frac{1}{z-z_1} + \frac{1}{z-z_2} + \dots + \frac{1}{z-z_n},$$

where z_1, z_2, \dots, z_n are the roots of $f(z)$. It follows from the form of (3) that we can interpret the conjugate imaginary complex quantity of the right-hand member of (3) as the force due to particles situated at the respective points z_1, z_2, \dots, z_n , each of which repels with a force equal to the reciprocal of the distance. The roots of (3) are the positions of equilibrium in the field of force due to these particles and an additional constant force of magnitude $1/c_1$ at every point of the plane. It is to be noted that a multiple root of $f(z)$ is always a root of $F_1(z)$ although not a root of (3).

The proof of Theorem II can be given so as to be almost identical with the proof of the limiting case for $c_1 = \infty$ of Theorem II.* That is, we consider one particular root z of $F_1(z)$, and for that point z we replace the n repelling particles of (3) by n equivalent particles which coincide at a point ζ . When the n particles have the circles C_i as their respective loci, the n -fold particle ζ has a

* This remark applies not to the proof of I, Theorem X for $k = 1$ as actually given in I, but to the proof there given in detail for I, Theorem VI. That proof can of course be given so as to apply directly to I, Theorem X, for $k = 1$.

circular region as its locus. Study of the locus of ζ will give us the theorem. Let us proceed to the general outline of the proof. We shall omit details if they are similar to the details given in I.

Theorem II is surely true when there is only one circle C_i . In this case, if a point z is on the boundary of the locus of the roots of $F_1(z)$ and is not on or within the circle C_1 which contains more than one particle, then it is necessary that

$$z_1 = z_2 = \dots = z_n,$$

and these points lie on C_1 .* Thus the line zz_1 is parallel to the axis of reals.

If the locus of all the roots of $f(z)$ is the *exterior* of a circle C_1 , the locus of the roots of $F_1(z)$ consists of the exterior of C_1 (if $n > 1$), and the exterior of another circle C'_1 defined as the circle traced by that root of $F_1(z)$ other than z_1 when the point $z_1 = z_2 = \dots = z_n$ traces the circle C_1 . If z is a root of $F_1(z)$ interior to the circle C_1 and on the boundary of its locus, z must lie on C'_1 , the point $z_1 = z_2 = \dots = z_n$ lies on C_1 , and the line zz_1 is parallel to the axis of reals.

We return to Theorem II for the polynomial $F_1(z)$, and suppose that there are precisely two circles C_i , namely C_1 and C_2 . It is clear from the field of force that no point z not lying on or between the common external tangents to C_1 and C_2 can be a root of (3). Moreover it follows from the special case of Theorem I mentioned above that not every point z lying between these tangents can be a point of the locus of the roots of (3). In the general case $n_1 > 1$, $n_2 > 1$, there are four circles C'_j of which two are the circles C_1 and C_2 , and one of which has its center interior to the interval α_1, α_2 .

* Here and below we make use of the lemma: *The force at a point P due to k particles situated in a circular region not containing P is equivalent to the force at P due to k coincident particles also in C.* See Walsh, TRANSACTIONS OF THIS SOCIETY, vol. 22 (1921), pp. 101-116, Lemma I, p. 102.

Let z be a fixed point on the boundary of the locus of the roots of $F_1(z)$. Then z is not interior to a circle C_1 or C_2 which is the locus of more than one root of $f(z)$, and is not interior to both circles, for in either of these cases z as well as any neighboring point can be made a multiple root of $f(z)$ and hence a root of $F_1(z)$. Suppose also that z is not on C_1 if $n_1 > 1$ nor on C_2 if $n_2 > 1$. Then (as in I, p. 47) we may replace the n_1 particles of C_1 by n_1 equivalent but coincident particles ζ_1 in C_1 , the n_2 particles of C_2 by n_2 equivalent but coincident particles ζ_2 in C_2 ; and finally we may replace these particles ζ_1 and ζ_2 by an equivalent n -fold particle ζ . When ζ_1 and ζ_2 have the circles C_1 and C_2 as their loci, the locus of ζ will be either the interior of a circle C or the exterior of a circle C , and C will cut the line zz' through z parallel to the axis of reals either at the same angle as do the circles C_1 and C_2 or at the supplementary angle according as the locus of ζ is interior or exterior to C .*

Corresponding to the particular point z that we are considering on the boundary of its locus, we may choose definite points $z_1, z_2, \dots, z_n, \zeta_1, \zeta_2, \zeta$, so that z is a position of equilibrium in the field of force. The point ζ must lie on the boundary of its locus C , and z must lie on the boundary of the locus of the roots of $F_1(z)$ when the locus of ζ ($= z_1 = z_2 = \dots = z_n$) is considered merely to be the locus C . For a small change in z makes only a small change in C . If z, ζ , and C have not the relation stated, a sufficiently small but arbitrary change of z can be made, and ζ can be chosen in the new locus C so that z is still a root of $F_1(z)$. That is, the original point z is not on the boundary of its locus.

The point ζ must therefore lie at an intersection of the

* It is conceivable that this locus of ζ should consist of the entire plane. This slight difficulty can be overcome as in I, p. 50. It is possible that the locus of ζ should be a half-plane. The reader will make the necessary modifications in the present treatment, which are indeed modifications of phraseology rather than of method.

line zz' and the circle C ; ζ is on the boundary of its locus. By the properties of the circles C_1 , C_2 , C , and the angles in which they are cut by the line zz' (as in I, p. 49), the points ζ_1 and ζ_2 must lie on C_1 and C_2 respectively and on the line zz' . In fact, ζ_1 and ζ_2 must be those intersections of zz' with C_1 and C_2 which cause z , the root of the corresponding polynomial $F_1(z)$, to lie on one of the circles C'_j . Thus we have proved that whenever z lies on the boundary of its locus, it lies on one of the circles C'_j .

The locus of the roots of $F_1(z)$ contains all the points interior to the circles C'_j ; indeed, it is obvious by a simultaneous translation of the points α_i as the roots of $f(z)$ and of the roots of the corresponding polynomial $F_1(z)$ that every point on or within a circle C'_j is a point of the locus. Every point of the boundary of the locus is a point of one of the circles C'_j , no point not between the two common external tangents of the circles C_i , C'_j is a point of the locus, so the locus consists of precisely the points stated in Theorem II.

The remark in Theorem II concerning the number of roots of $F_1(z)$ in the regions C'_j follows (as in I, p. 50) from a consideration of the roots of $F_1(z)$ when the roots of $f(z)$ are the points α_i , and from the continuity of the roots of $F_1(z)$ considered as functions of the roots of $f(z)$. Theorem II is now completely proved for the case $k = 2$.

The same method is used in every case. We replace the particles z_1, z_2, \dots, z_n by a single equivalent particle ζ , whose locus is a circular region. This particle ζ is on the boundary of its locus if z is on the boundary of its locus, and the properties of the boundary of the locus of ζ in connection with the properties of the circles C_i enable us to prove that the point z is on one of the circles C'_j , and to prove the theorem in its generality. Further details are left to the reader.

Theorem II is a result dealing with the interiors of circles as the loci of roots of polynomials. We may use the same method to prove a result similarly dealing with the *exteriors* of equal circles whose centers lie on a line parallel to the

$F_2(z), F_3(z), \dots, F_{n-1}(z)$. This condition will always be satisfied in Theorem I if, when the polynomial (1) and the points α_i are given, the common radius of the circles C_i is sufficiently small, and this is the meaning we give to the restriction in the enunciation of the theorem, "if the circles C_i are sufficiently small". In order to prove the theorem up to but not including the last sentence, we need not require that no two of the circles C'_j which are the loci of the roots of $F_n(z) = F(z)$ should have a point in common, but if this condition is satisfied the number of roots of $F(z)$ in those circles C'_j is as indicated. This fact is proved (as in I, p. 50) from the continuity of the roots of $F(z)$ considered as functions of the roots of $f(z)$.

It still remains, in the proof of Theorem I, to remove the restriction $a_0 \neq 0$. This can be accomplished by remembering that Theorem I has already been proved* when $F(z)$ is simply a derivative of $f(z)$. Hence application of Theorem I for the case $a_0 \neq 0$ to a derivative of $f(z)$ shows that the roots of $F(z)$ must lie in the circles C'_j . Every point on or within the circles C'_j is a point of the locus, for this follows by a translation. We naturally require as before that the circles C_i be sufficiently small. The number of roots of $F(z)$ in a circle C'_j can be determined by continuity, so Theorem I is completely established.

We add the statement of a theorem which is almost trivial when the preceding development is considered, but which seems nowhere to have been mentioned in the literature.†

THEOREM III. *If the roots of a polynomial $f(z)$ of degree n lie in a region R bounded by two parallels to the axis of reals, and if the roots of (1) are all real, then the roots of*

* In I, Theorem X. The proof of the present note is also valid for this case; we merely choose the left-hand member of (3) to be zero to prove the result for the first derivative, and apply repeatedly for the other derivatives.

† In connection with these theorems for the polynomial $F_1(z)$, see Fujiwara, *TÔHOKU MATHEMATICAL JOURNAL*, vol. 9 (1916), pp. 102-108. See also Uchida, *TÔHOKU MATHEMATICAL JOURNAL*, vol. 10 (1916), pp. 139-141.

$$F(z) = a_0 f(z) + a_1 f'(z) + a_2 f''(z) + \dots + a_n f^{(n)}(z)$$

also lie in R .

The generalization of Theorem I for the case $k = 1$ but which makes no restriction on (1) nor on the size of the circles C_i (i. e., II, Theorem VI) gives us the following theorem.

THEOREM IV. *If all the roots of a polynomial $f(z)$ lie in the strip*

$$a_1 \leq y \leq a_2$$

of the complex ($z = x + iy$) plane, and if all the roots of the polynomial (1) lie in the strip

$$b_1 \leq y \leq b_2,$$

then all the roots of

$$F(z) = a_0 f(z) + a_1 f'(z) + a_2 f''(z) + \dots + a_n f^{(n)}(z)$$

lie in the strip

$$a_1 + b_1 \leq y \leq a_2 + b_2.$$

Innumerable theorems of the same nature as Theorem IV can be proved by II, Theorem VI. We give one further example, after a preliminary remark.

If two convex regions R_1 and R_2 are the respective loci of points z_1 and z_2 , then the region R which is the locus of the point $z = z_1 + z_2$ is also a convex region. We shall speak of R as the *sum* of the regions R_1 and R_2 . Then we have

THEOREM V. *If two convex regions are the respective loci of the roots of $f(z)$ and of (1), then their sum is the locus of the roots of $F(z)$.*

Theorem V obviously includes Lucas's theorem for convex regions on the roots of the derivative of a polynomial. Theorem V includes, and can be proved from Theorem IV. Theorem IV includes Theorem III, and Theorem III includes the theorem due to Hermite which we have been using in Theorem I, to the effect that if the roots of $f(z)$ and of (1) are real, then the roots of $F(z)$ are also real.*

* Reference has been made Takagi's work in this connection.

Theorem V does not directly extend to the case of infinite convex

Let us study the case of real polynomials in more detail. We shall prove for them a theorem analogous to Theorem I.

THEOREM VI. *Let intervals I_i ($i = 1, 2, \dots, k$) of the axis of reals, whose extremities are α_i, β_i ($\alpha_i \leq \beta_i$) be the respective loci of n_i roots of a variable polynomial $f(z)$ which has no other roots. Then if the roots of the polynomial (1) are all real, the locus of the roots of $F(z)$ is composed of a number of intervals I'_j of the axis of reals. The left-hand extremities of the intervals I'_j are the roots of $F(z)$ when the roots of $f(z)$ are concentrated at the points α_i ; the right-hand extremities are the corresponding roots of $F(z)$ when the roots of $f(z)$ are concentrated at the points β_i . Any interval I'_j which has no point in common with any other interval I'_j contains a number of roots of $F(z)$ equal to the multiplicity of its left-hand extremity as a root of $F(z)$ when the roots of $f(z)$ are the points α_i . If the intervals I_i are all of the same length, the intervals I'_j are also all of this same length.*

Theorem VI has been proved (II, footnote, p. 180) for the case $k = 1$, and has been proved (I, Theorem XV) for general k for the case that $F(z)$ is a derivative of $f(z)$. The proof of this latter theorem is to be followed closely in the proof of Theorem VI. We shall prove Theorem VI under the assumption that none of the intervals I_i is a point; to include this more general case requires merely a slight change in phraseology. In the theorem as stated the intervals are assumed to be finite, but the theorem can be extended to include infinite intervals. We prove the theorem first for the case of the polynomial

$$F_1(z) = f(z) + c_1 f'(z).$$

regions, using the ordinary extension of the definition of convexity, although the following is an immediate result of II, Theorem VI:

If the roots of (1) have as their locus any region R and if the roots of $f(z)$ have as their locus the region C ,

$$|z - \alpha| \leq r, \text{ where } \alpha \text{ and } r \text{ are fixed,}$$

then the locus of the roots of $F(z)$ is the region S which is the locus of points

$$\zeta = \rho + z,$$

where ρ and z have the respective loci R and C . If R is the interior or exterior of a circle, so also is S .

Let us denote by $\alpha_i^{(m)}$ the roots of $F_m(z)$ when the roots of $f(z)$ are concentrated at the points α_i , $\alpha_i^{(m)} \leq \alpha_j^{(m)}$ when $i < j$, and similarly by $\beta_i^{(m)}$ the roots of $F_m(z)$ when the roots of $f(z)$ are concentrated at the points β_i , $\beta_i^{(m)} \leq \beta_j^{(m)}$ when $i < j$. We shall show that the intervals $(\alpha_i^{(m)}, \beta_i^{(m)})$ form the locus of the roots of $F_m(z)$.

We start with the roots of $f(z)$ concentrated at the points α_i and move these roots continuously toward the right until they reach the points β_i . The roots of $F_1(z)$ also vary continuously; they start at the points α_i' and reach the points β_i' . We can even say that the r th root z_r' of $F_1(z)$ (roots numbered in order from the left) varies continuously. Let us now prove that z_r' always moves, if at all, toward the right.

The equation determining z_r' is of the form

$$(6) \quad \Phi = \frac{1}{c_1} + \frac{f'(z)}{f(z)} = \frac{1}{c_1} + \frac{m_1}{z_r' - \gamma_1} + \frac{m_2}{z_r' - \gamma_2} + \dots + \frac{m_s}{z_r' - \gamma_s},$$

where the γ_i are the roots of $f(z)$, coinciding in any multiplicities desired. We compute the values

$$\frac{\partial \Phi}{\partial z_r'} = -\frac{m_1}{(z_r' - \gamma_1)^2} - \frac{m_2}{(z_r' - \gamma_2)^2} - \dots - \frac{m_s}{(z_r' - \gamma_s)^2},$$

$$\frac{\partial \Phi}{\partial \gamma_i} = \frac{m_i}{(z_r' - \gamma_i)^2}.$$

It is always true that $\partial z_r' / \partial \gamma_i$ is positive, so z_r' always increases with γ_i .*

* This result is the chief tool in the proof of Theorem VI, and as it is true that $\partial z_r' / \partial \gamma_i \neq 0$ for a large class of polynomials, Theorem VI can be suitably modified to apply to much more general polynomials, in particular to $f'(z)$, where $f(z)$ is real but has not necessarily all real roots. Theorem VI holds without change for polynomials and the iterates of polynomials of the type

$$(a - b^2z)f(z) + cf'(z),$$

where a, b, c are real and the roots of $f(z)$ are real. This last class of polynomials has recently been considered by Nagy, *JAHRESBERICHT DER VEREINIGUNG*, vol. 31 (1922), pp. 238-251.

Equation (6) is no longer valid to determine z'_r if z'_r is located at a multiple root of $f(z)$. Under these circumstances, if γ_i does not coincide with z'_r , the motion of γ_i does not change the position of z'_r . If γ_i does coincide with z'_r and if γ_i is moved to the right, z'_r is either unchanged or moved to the right; this follows immediately from the fact that a t -fold root of $f(z)$ is a $(t-1)$ -fold root of $F_1(z)$, and from the fact that every interval of the axis of reals bounded by roots of $f(z)$ contains at least one root of $F_1(z)$.

From the general fact, then, that the r th root z'_r of $F_1(z)$ varies continuously and in one sense under the indicated variation of the roots of $f(z)$, it follows that z'_r traces the entire interval from α'_r to β'_r and can never lie outside that interval. The determination of the locus of Theorem VI is now complete for the polynomial $F_1(z)$. The number of roots of $F_1(z)$ in an interval (α'_r, β'_r) having no point in common with any other interval (α'_s, β'_s) is easily found by continuity.

When the roots of $f(z)$ move from the points α_i continuously to the right, the roots of $F_1(z)$ move from the points α'_i continuously to the right, and from the formulas (5), the roots of $F_2(z)$ move from the points α''_i continuously to the right, and the roots of $F_m(z)$ move from the points $\alpha_i^{(m)}$ continuously and to the right. This gives us Theorem VI for the case $a_0 \neq 0$; the restriction $a_0 \neq 0$ can be removed with no difficulty as in the proof Theorem I of the present paper.

HARVARD UNIVERSITY