

ON THE LOCATION OF THE ROOTS OF CERTAIN TYPES OF POLYNOMIALS*

BY

J. L. WALSH

When we study the dependence of a variable on k other variables which vary independently, our problem may be very much simplified if we can consider all or some of these independent variables to coincide and thus study the dependence of our original dependent variable on one new variable or at least on a number of new independent variables less than k . The present writer has recently published a theorem (Theorem II, below) which enables us to make a reduction of this sort in the study of the relations between the roots of certain types of polynomials. The present paper aims to prove Theorem I (below), which is a much more general result of the same nature, and to indicate various applications of that theorem. The applications given are extremely simple and follow from Theorem I with practically no further machinery.† The most interesting application is Theorem VI.

Our problem is, more explicitly, to study the geometric relationship of the roots of a polynomial

$$f(z) = (z - a_1)(z - a_2) \cdots (z - a_m)$$

to the roots of a related polynomial

$$\phi(z) = (z - b_1)(z - b_2) \cdots (z - b_n)$$

* Presented to the Society, December 29, 1920, and April 23, 1921.

† When this paper was first offered for publication, the writer believed Theorem I to be new. Professor D. R. Curtiss has kindly pointed out its connection with a theorem due to Grace and has indicated an entirely new proof of Theorem I; the reader will refer to Professor Curtiss's note which immediately follows this paper.

The *point of view* of the present paper in the proof and application of Theorem I seems to be new, and also the results obtained except where otherwise stated.

This paper is the development of a short note published in the *Paris Comptes Rendus*, March, 1921, to which explicit reference is made later, and which contained in outline the proof of Theorem I. In the interval between the publication of that note and the publication of the present paper, there have appeared a number of other papers dealing with Grace's Theorem. See Szegő, *Mathematische Zeitschrift*, vol. 13 (1922), pp. 28-55; Cohn, *Mathematische Zeitschrift*, vol. 14 (1922), pp. 110-148; Egerváry, *Acta Litterarum ac Scientiarum, Regiae Universitatis Hungaricae Francisco-Josephinae*, vol. 1 (1922), pp. 39-45; Fekete, same Journal, vol. 1 (1923), pp. 98-100.

The present paper has very little in common with these other papers, but with Szegő's Theorem 9 the reader should compare our Theorem X, and with Szegő's Theorems 13-15 compare our Theorem IV and also Walsh, *American Mathematical Monthly*, vol. 29 (1922), pp. 112-114.

whose roots are supposed to be known.* When we study the functional dependence of a root of f , say a_1 for definiteness, on the b_i , it may occur that without changing a_1 we can replace the n roots b_i by n roots of a polynomial ϕ which coincide at some point b which bears a simple relation to the roots b_i . Then we can study the dependence of a_1 on b instead of on b_1, b_2, \dots, b_n , which is frequently a simplification of our problem.

To the plane of the complex variable we adjoin the point at infinity; infinity is to be considered simply as an ordinary value of the variable. As a geometrical consequence we shall consider the term *circle* to include the possibility of a straight line. We shall have occasion to deal with *circular regions*, by which we mean a closed portion of the plane bounded by a single circle, that is, the interior of a circle, the exterior of a circle (including the point of infinity), a half-plane, a point, or the entire plane; the points of the boundary are always to be included in the region.

I. A GENERAL THEOREM

We proceed to the proof of our main result:

THEOREM I. *Let $f(z)$ be a polynomial in z whose coefficients are polynomials linear in and symmetric in each of the sets of variables*

$$\{\alpha_1, \alpha_2, \dots, \alpha_k\}, \quad \{\beta_1, \beta_2, \dots, \beta_l\}, \quad \dots, \quad \{\lambda_1, \lambda_2, \dots, \lambda_g\}.\dagger$$

Let these points $\{\alpha_i\}$, $\{\beta_i\}$, \dots , $\{\lambda_i\}$ lie in circular regions $C_\alpha, C_\beta, \dots, C_\lambda$. Then for any fixed values of these variables and of z we can always make all the $\{\alpha_i\}$ coincide in C_α , all the $\{\beta_i\}$ coincide in C_β , etc., without altering the value of $f(z)$.

The theorem also obtains if we replace $f(z)$ by the quotient of two polynomials of the type described, except that we are to consider the conclusion of the theorem satisfied if these two polynomials vanish simultaneously.‡ Of course if we have two polynomials of the type required, their sum is also of that type.

We shall prove the theorem considering the $\{\beta_i\}$, $\{\gamma_i\}$, \dots , $\{\lambda_i\}$ always as fixed and showing that we can make the $\{\alpha_i\}$ coincide in C_α as stated. Then we can consider the $\{\alpha_i\}$ fixed and coincident, and the $\{\gamma_i\}$, \dots , $\{\lambda_i\}$ fixed; our former reasoning will show that the $\{\beta_i\}$ can be made to coincide. Continued reasoning in this manner will evidently complete the proof of the theorem. It will be convenient to assume that the value of $f(z)$ considered

* Professor Curtiss has recently published a very interesting report on this general field, *Science*, vol. 55 (1922), pp. 189-194.

† The coefficients of $f(z)$ need not be homogeneous in each of these sets of variables, but each coefficient must be a linear combination of the elementary symmetric functions of each of these sets with coefficients linear combinations of the elementary symmetric functions of the other sets. These linear combinations may, moreover, contain constant terms.

‡ This is what actually occurs in the situation of Theorem II if we choose P inside C .

in the hypothesis is zero; this involves no loss of generality, for the addition of a constant term to $f(z)$ does not alter the properties required.

We now wish to show that the α_i can all be made to coincide. If the region C_α is a point, the statement is trivial. If we prove the theorem where C_α is a circular region not the whole plane we have proved it where C_α is the whole plane.

Consider the α_i to vary independently and to have the region C_α as their common locus. Then the relation

$$(1) \quad f(z) = 0$$

defines z as an analytic function of the α_i and hence z will have a certain locus Z ; this locus Z will be a closed point set since C_α is closed. If (1) degenerates and does not effectively contain z at all, we introduce an auxiliary variable ζ by placing

$$F(\zeta) \equiv f(z) - \zeta = 0;$$

the new function $F(\zeta)$ surely contains ζ and we may reason with it as before for $f(z)$.

Let α be any point interior to C_α . There is a certain locus Z of points z corresponding to the null-circle α as the locus of the $\{\alpha_i\}$. Make use of the auxiliary circle C of the coaxial family determined by α and C_α . Let C commence with the position α and gradually enlarge and coincide with C_α , the region bounded by C and containing α always considered as the locus of the $\{\alpha_i\}$. The locus Z also grows larger and varies continuously with C (except as noted below), for the roots z of (1) are continuous functions of the $\{\alpha_i\}$. When the region C continues to enlarge, beyond C_α if necessary, the locus Z eventually becomes the entire plane; this happens ordinarily before C has swept through the entire plane. The point z with which we start is a point of the locus Z surely when C coincides with C_α and possibly before.

The case that the locus Z does not vary continuously with C occurs only if, for some choice of the $\{\alpha_i\}$, (1) vanishes identically in z . When this occurs, the roots $\{z\}$ of (1) do not vary continuously with the $\{\alpha_i\}$. The locus Z may enlarge suddenly and become the entire plane while its boundary does not necessarily sweep over the whole plane.

Whether or not this phenomenon takes place, corresponding to our original point z there is some circle C such that for no circle C' , the region C' smaller than and entirely contained in the region C , can z be a point of the locus Z of the roots of (1). This statement and in fact Theorem I as well are true if z is a point of the locus corresponding to α . The statement is true in any other case. For if there is a sequence of sets of points $\{\alpha_i\}$ in a sequence of regions each contained in the preceding, there is at least one limiting set of points $\{\alpha_i\}$ contained in all the regions and hence in the limit region. The relation (1) obtains for this limit set.

Our entire proof of Theorem I rests on the remark that when we fix z and all but two of the points $\{\alpha_i\}$, equation (1) becomes a homographic and involutory relation between the other two of these points and hence when one of these points traces a circle the other also traces a circle. The proof is complicated, however, by the possibility that this relation may degenerate and may effectively contain but one or neither of these variables.*

We prove now a very special case of Theorem I, namely, that if a circular region C contains two points α_1 and α_2 connected by a relation of the form

$$(2) \quad a_1 \alpha_1 \alpha_2 + a_2 (\alpha_1 + \alpha_2) + a_3 = 0,$$

then C contains a root α' of the equation

$$(3) \quad a_1 \alpha'^2 + 2a_2 \alpha' + a_3 = 0.$$

Transform (2) by an auxiliary linear transformation so that the point at infinity is a double point of the transformation (α_1, α_2) defined by (2). The line through α_1 and α_2 is transformed into itself by the transformation (α_1, α_2) , for the three points $(\alpha_1, \alpha_2, \infty)$ go into $(\alpha_2, \alpha_1, \infty)$. When α_1 is moved on this line, α_2 moves on the line in the opposite sense, for if the two points move always in the same sense, when α_1 is moved along the finite segment from its original position to the original position of α_2 , α_2 has moved in the same sense, has at no time coincided with the point at infinity and hence has not reached the original position of α_1 . The points α_1 and α_2 moving on the line joining them in opposite senses must have a double point D between them. It is thus evident that the region C which contains α_1 and α_2 must contain D or the point at infinity and hence a solution of (3).

The reasoning just used supposes implicitly that the relation (2) does not degenerate. If (2) does not contain α_2 for a particular value of α_1 , that value of α_1 is a solution of (3). If (2) is satisfied identically, any values α_1 and α_2 suffice.

We now return to the proof of Theorem I in its general form. Let C be the

* Thus we may start with the auxiliary polynomial

$$F(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_k) \quad k > 3$$

and consider for the polynomial of Theorem I

$$f(z) \equiv F'(z).$$

The relation

$$(1') \quad f(z) = 0$$

if $\alpha_{k-1} = \alpha_k = z$ contains effectively none of the variables $\alpha_1, \alpha_2, \dots, \alpha_{k-2}$. If $\alpha_k = z$, we have (1') reducing to

$$(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_{k-1}) = 0,$$

and hence we have no effective homographic relationship between α_{k-1} and α_k .

This example is not an unnatural one; indeed it is one of the first to which we should think of applying Theorem I; compare Theorem II and its applications to which a reference is given.

smallest region of the kind already described for which z is a point of the locus Z . There is at least one of the points α_i on C . We shall show that all those points can be made to coincide on C .

Two points α_1 and α_2 which are on C can be made to coincide either on C or within C . For we fix z and the remainder of the points $\{\alpha_i\}$ and hence have a relation of type (2) between α_1 and α_2 . A double point of the transformation is either on the circle C or interior to the region C .

Let us combine in this manner as many of the points $\{\alpha_i\}$ on C as possible, so as to leave the smallest number of points $\{\alpha_i\}$ actually on C . This number of points we denote by n ; it is greater than zero and cannot be reduced.

We shall prove by induction that all these n points on C can be made to coalesce on C . This fact is evidently true for two points, for no double point of the corresponding transformation can be interior to C . We assume explicitly that whenever there are given $n - 1$ points of this sort on an arc A of C , they can be made to coincide on A without changing any other point α_i , or the relation (1); we shall prove this fact for n points. It involves no loss of generality to suppose $n - 1$ of these points at an end point X of A and the n th point at the other end point Y .

We consider one of the points α_1 at X and the point α_n at Y as connected by (1), while z and the other $\{\alpha_i\}$ remain fixed. When α_1 moves on C , α_n moves on C in the opposite sense, and the transformation (α_1, α_n) has a double point D interior to A . Consider α_1 and α_n to coincide at D . The new arc bounded by X and D contains the n points. We can make coincide, at a point X' , $n - 1$ points α_i , X' on the arc XD , and the other point α_n will be at D . Then we have the n points on an arc A' bounded by X' and Y' , the arc A' shorter than A and every point of A' a point of A . Moreover, there are $n - 1$ points at X' , the end point nearest X , and the other is at $Y' = D$, the end point nearest Y .

We can continue in this manner successively to shorten our arc A , and we can shorten it indefinitely. For let us suppose we have a sequence of arcs (using the natural extension of our former notation), $XY, X'Y', X''Y'', \dots$, and suppose the points X, X', X'', \dots have a limit point x and the points Y, Y', Y'', \dots a limit point y different from x and of such a nature that we cannot make our arc of type A shorter than xy . We have $n - 1$ coincident points α_i , successively approaching x , and one point α_n simultaneously approaching y . From the continuity of the left-hand member of (1) we can therefore suppose these $n - 1$ points α_i to coincide at x and the other point α_n to lie at y . Then the procedure formerly used shows that we can bring the n points into an arc shorter than xy all of whose points are points of xy . This contradiction completes the proof of our statement that the arcs A, A' ,

A'' , \dots shorten indefinitely, and hence the points $\alpha_1, \alpha_2, \dots, \alpha_n$ can all be made to coincide at the limit point P of those arcs, with (1) still satisfied.

If $k = n$, Theorem I is proved. If $k > n$, consider the transformation (α_1, α_k) defined as before. We have supposed that none of the points $\alpha_1, \alpha_2, \dots, \alpha_n$ can be moved from the circle C , so the transformation cannot contain α_k and hence α_k can be moved as near P as desired. Thus any other of the points $\alpha_{n+1}, \alpha_{n+2}, \dots, \alpha_k$ can be moved as near P as desired without changing the value of $f(z)$. Hence all these points can be made to approach P and so the value of $f(z)$ is unchanged if all the $\alpha_1, \alpha_2, \dots, \alpha_k$ coincide at P . Theorem I is now completely proved.

II. SPECIAL CASES OF THEOREM I, WITH APPLICATIONS

There are two results, essentially special cases of Theorem I, which are particularly interesting in their applications.

THEOREM II. *If the points $\alpha_1, \alpha_2, \dots, \alpha_k$ lie in a circular region C and if z is exterior to C , the root of the equation in α*

$$\frac{1}{z - \alpha_1} + \frac{1}{z - \alpha_2} + \dots + \frac{1}{z - \alpha_k} = \frac{k}{z - \alpha}$$

*lies in C .**

Theorem II is in reality a special case of the extension of Theorem I where we consider the quotient of two polynomials of the kind described. The denominator polynomial cannot vanish, since z is exterior to C .

For applications of Theorem II we refer to the citation already made and that made in the next following footnote. We proceed to another special case of Theorem I:

THEOREM III. *If the points $\alpha_1, \alpha_2, \dots, \alpha_k$ lie in a circular region C , the equation in α*

$$(4) \quad (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_k) = (z - \alpha)^k$$

has at least one root in C .

Theorem III can be proved independently of Theorem I in a manner precisely analogous to the proof of Theorem II (*loc. cit.*). This proof of Theorem III involves a transcendental transformation of the α - (or (x, y) -) plane:

$$\alpha - z = x + iy = e^{u+iv},$$

and a study of the transform of C and certain centers of gravity in the (u, v) -plane. This proof gives in certain cases more detailed information than does Theorem I regarding the root α of (4). Thus if z is exterior to C , which is

* Walsh, these *Transactions*, vol. 22 (1921), p. 102; Lemma I.

Theorem II is closely connected with another more simple corollary of Theorem I, namely, that if k equal particles lie in a circle their center of gravity also lies in that circle.

the interior of a circle, we can write

$$k \log (z - \alpha) = \log (z - \alpha_1) + \log (z - \alpha_2) + \cdots + \log (z - \alpha_k),$$

where all the logarithms involved have the coefficients of $\sqrt{-1}$ in their pure imaginary parts lying between two numbers whose difference is less than π . We add the remark, without proof at this time, that there is no other finite region bounded by a regular curve which possesses the property of the finite circular region indicated by Theorem III.*

Theorem III can easily be extended to give some information concerning the location of *all* the roots of equation (4) in α . In fact, if α' is any root, the other roots are given by

$$\begin{aligned} (z - \alpha)^k &= (z - \alpha')^k, \\ z - \alpha &= \omega (z - \alpha'), \end{aligned}$$

where ω is a k th root of unity. If we suppose that α' is in C , all the roots of (4) must lie in the k circular regions obtained by revolving C about z as center of rotation through the angles $2j\pi/k, j = 1, 2, \dots, k$. In particular suppose one of these k circular regions, say C_1 , is external to all the other $k - 1$ circular regions. Then we can prove that C_1 has on or within it precisely one root of (4). For consider the points $\alpha_1, \alpha_2, \dots, \alpha_k$ to move continuously always remaining in C and to coincide at the center of C . In this situation C_1 contains precisely one root of (4). The roots of (4) vary continuously with the points $\alpha_1, \alpha_2, \dots, \alpha_k$; none can enter or leave C_1 during motion of the kind indicated; so in the original situation C_1 contained precisely one root of (4).

However, our purpose is not primarily to study the roots of (4), but rather to use Theorems I and III as tools in proving more interesting relations. In preparation for these results we now prove the

LEMMA. *If the interiors and boundaries of the two circles C_1 and C_2 , whose centers are α_1 and α_2 and radii r_1 and r_2 respectively, are the loci of two points z_1 and z_2 , then as z_1 and z_2 vary independently, the locus of the point z which divides the segment (z_1, z_2) in the constant ratio $(m_1 : m_2)$,*

$$(5) \quad z = \frac{m_2 z_1 + m_1 z_2}{m_1 + m_2} \quad (m_1 \neq -m_2, \quad m_1 m_2 \neq 0),$$

is the interior and boundary of a circle C whose center is

$$\frac{m_2 \alpha_1 + m_1 \alpha_2}{m_1 + m_2}$$

and whose radius is

$$\left| \frac{m_2 r_1}{m_1 + m_2} \right| + \left| \frac{m_1 r_2}{m_1 + m_2} \right|. \dagger$$

* For the corresponding fact for Theorem II, compare Walsh, these Transactions, vol. 24 (1922), pp. 31-69; Theorem III.

† The terminology that z divides the segment (z_1, z_2) in the ratio $(m_1 : m_2)$ is usual when

If z_1 and z_2 are on or within C_1 and C_2 respectively,

$$|z_1 - \alpha_1| \leq r_1, \quad |z_2 - \alpha_2| \leq r_2,$$

we have

$$z - \frac{m_2 \alpha_1 + m_1 \alpha_2}{m_1 + m_2} = \frac{m_2}{m_1 + m_2} (z_1 - \alpha_1) + \frac{m_1}{m_1 + m_2} (z_2 - \alpha_2),$$

which is in absolute value less than or equal to

$$\left| \frac{m_2 r_1}{m_1 + m_2} \right| + \left| \frac{m_1 r_2}{m_1 + m_2} \right|.$$

It remains to be shown that given any point z on or within C , we can properly determine z_1 and z_2 . In order to do this, we merely place

$$\frac{m_2 (z_1 - \alpha_1)}{m_1 + m_2} = \left(z - \frac{m_2 \alpha_1 + m_1 \alpha_2}{m_1 + m_2} \right) \frac{\left| \frac{m_2 r_1}{m_1 + m_2} \right|}{\left| \frac{m_2 r_1}{m_1 + m_2} \right| + \left| \frac{m_1 r_2}{m_1 + m_2} \right|},$$

$$\frac{m_1 (z_2 - \alpha_2)}{m_1 + m_2} = \left(z - \frac{m_2 \alpha_1 + m_1 \alpha_2}{m_1 + m_2} \right) \frac{\left| \frac{m_1 r_2}{m_1 + m_2} \right|}{\left| \frac{m_2 r_1}{m_1 + m_2} \right| + \left| \frac{m_1 r_2}{m_1 + m_2} \right|}.$$

Then z_1 and z_2 lie in the proper regions and satisfy (5), so the proof of the lemma is complete.

This lemma gives no result if $m_1 + m_2 = 0$; to treat this case we take (5) in the form

$$(6) \quad m_2 (z - z_1) = -m_1 (z - z_2),$$

which is equivalent for our present purpose. If C_1 and C_2 are mutually external there is no point z different from the point at infinity which satisfies (6). If C_1 and C_2 are not mutually external we may choose $z_1 = z_2$, and every point of the plane satisfies (6). In the theorems below we give the general formulas for the determination of C with the understanding that when $m_1 + m_2 = 0$, C is considered to contain no finite point of the plane or every finite point of the plane according as C_1 and C_2 are or are not mutually external.*

m_1 and m_2 are real; whether m_1 and m_2 are real or not, we simply understand that statement to mean that z is given by (5).

This lemma was proved by the present writer and by the present method for m_1 and m_2 real and positive, in a note in the *Comptes Rendus du Congrès international*, Strasbourg, 1920. See also Theorem II of the reference given in connection with the present Theorem II.

* Thus, a careful statement of Theorem IV for the case $A = 1$ is

THEOREM IVa. *If the points a_1, a_2, \dots, a_k have as their locus a circle C_1 whose center is α and radius r_1 , and if the points b_1, b_2, \dots, b_k have as their locus a circle C_2 whose center is β*

We now apply Theorem III and the Lemma to prove

THEOREM IV. *If the points a_1, a_2, \dots, a_k have as their locus (the interior and boundary of) a circle C_1 whose center is α and radius r_1 , and if the points b_1, b_2, \dots, b_k have as their locus a circle C_2 whose center is β and radius r_2 , the roots of the equation*

$$(7) \quad (z - a_1)(z - a_2) \cdots (z - a_k) - A(z - b_1)(z - b_2) \cdots (z - b_k) = 0$$

have as their locus the k circles with the respective radii

$$\frac{r_1 + |A^{1/k}|r_2}{|1 - A^{1/k}|}$$

and centers

$$(8) \quad \frac{\alpha - A^{1/k}\beta}{1 - A^{1/k}}$$

where $A^{1/k}$ takes all the k values possible. Any one of these k circles which is external to all the others contains precisely one root of (7).

It is evident that any polynomial equation can be written in the form (7), and in an infinite variety of ways.

We prove Theorem IV as follows. If a point z is a root of (7), an application of Theorem III shows that

$$(z - a)^k - A(z - b)^k = 0,$$

where a and b are some points on or within C_1 and C_2 respectively. Hence we have

$$z - a = A^{1/k}(z - b),$$

$$z = \frac{a - A^{1/k}b}{1 - A^{1/k}},$$

where $A^{1/k}$ is some k th root of A , so z is on or within one of the k circles of the theorem. Variation of a and b independently and over C_1 and C_2 as loci shows that the k circles are the actual loci of the roots of (7). Continuous motion of the points $a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_k$ so as to remain in their proper regions and radius r_2 , and if C_1 and C_2 are mutually external, then the roots of the equation

$$(7') \quad (z - a_1)(z - a_2) \cdots (z - a_k) = (z - b_1)(z - b_2) \cdots (z - b_k)$$

have as their locus the $k - 1$ circles with the respective radii

$$\frac{r_1 + r_2}{|1 - \omega|}$$

and centers

$$\frac{\alpha - \omega\beta}{1 - \omega}$$

where ω takes the $k - 1$ values of the k th roots of unity which differ from unity itself. Any one of these $k - 1$ circles which is external to all the others contains precisely one root of (7').

If C_1 and C_2 are not mutually external, the locus of the roots of (7') is the entire plane.

The details of the proof of Theorem IVa are similar to those of the proof of Theorem IV and are left to the reader.

and finally to coincide shows in precisely the manner used in considering the roots of (4) that any one of the k circles which is exterior to all the others contains precisely one root of (7).

Theorem IV reduces to the lemma when $k = 1$.

A few remarks on the geometry of the situation are not out of place. The k points (8) all lie on the circle which is the locus of points z such that

$$\left| \frac{z - \alpha}{z - \beta} \right| = |A^{1/k}|,$$

which circle is of course a circle of the coaxial family determined by α and β as null circles. Each of the k points also lies on a circle which is the locus of points z such that

$$\frac{\frac{z - \alpha}{|z - \alpha|}}{\frac{z - \beta}{|z - \beta|}} = \frac{A^{1/k}}{|A^{1/k}|};$$

that is, the argument of the ratio $(z - \alpha)/(z - \beta)$ is constant. These k circles belong to the coaxial family of all circles passing through α and β , the family conjugate to the former coaxial family mentioned. These same k circles can be arranged in order so that at α and β each cuts its predecessor at an angle of $2\pi/k$.

It is quite easy for us to obtain results concerning the roots of the derivatives of equation (7). The m th derivative ($k > m \geq 0$), except for a constant factor, can be written

$$(9) \quad (z - a'_1)(z - a'_2) \cdots (z - a'_{k-m}) - A(z - b'_1)(z - b'_2) \cdots (z - b'_{k-m}) = 0.$$

The points a'_j all lie on or within C_1 and the points b'_j all lie on or within C_2 , by the theorem of Lucas concerning the roots of the derivative of a polynomial. Hence all the roots of (9) lie on or within the $k - m$ circles with the respective radii

$$\frac{r_1 + |A^{1/(k-m)}| r_2}{|1 - A^{1/(k-m)}|}$$

and centers

$$\frac{\alpha - A^{1/(k-m)} \beta}{1 - A^{1/(k-m)}},$$

where $A^{1/(k-m)}$ takes all the $k - m$ values possible. These circles form the locus of the roots of (9). Any one of these $k - m$ circles which is external to the remaining $k - m - 1$ circles contains precisely one root of (9). This new result includes Theorem IV as the case $m = 0$.

We now state another theorem which results from Theorem III and which like Theorem IV refers to the sum of two polynomials.*

THEOREM V. *If the points a_1, a_2, \dots, a_k lie on or within a circle C whose center is α and radius r , all the roots of the equation*

$$(10) \quad (z - a_1)(z - a_2) \cdots (z - a_k) - A = 0$$

lie on or within one of the k circles which have the common radius r and whose centers are the points $\alpha + A^{1/k}$, where $A^{1/k}$ takes all the k possible values. If these k circles are mutually external, each contains precisely one root of (10).

Of course any polynomial equation can be written in the form (10) and in an infinite variety of ways. The proof of the theorem follows the proof of Theorem IV. If C is the locus of a_1, a_2, \dots, a_k , the k circles form the locus of the roots of (10).

Equation (10) is particularly interesting because the derived equation is independent of A . All the roots of the m th derived equation, $m < k$, lie on or within C . Theorem V can also be extended to a polynomial of the form

$$\frac{d}{dz} [(z - a_1)(z - a_2) \cdots (z - a_k)] - A = 0,$$

but this generalization as well as Theorem V itself is included in a more general theorem to be proved from Theorem I.

III. SOME DIRECT APPLICATIONS OF THEOREM I

We now proceed to derive a number of results directly from Theorem I instead of from Theorem II or Theorem III.

If in Theorem I the polynomial $f(z)$ contains merely one set of variables $\{\alpha_i\}$ referring to a single circular region, the locus Z can have as its boundary no point other than a point of the curve traced by z when $\alpha = \alpha_1 = \alpha_2 = \dots = \alpha_k$ traces the circle C_α . This leads to

THEOREM VI. *Let the common locus of all the roots of a polynomial*

$$f(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_k)$$

be the interior and boundary of a circle C whose center is α and radius r . Then the locus of the roots of the polynomial

$$(11) \quad A_0 f(z) + A_1 f'(z) + \cdots + A_{k-1} f^{(k-1)}(z) + A_k f^{(k)}(z)$$

is composed of the interiors and boundaries of the circles whose common radius is r and whose centers are the roots of

$$(12) \quad A_0 (z - \alpha)^k + kA_1 (z - \alpha)^{k-1} + \cdots + k(k-1) \cdots 2 \cdot 1A_k.$$

* If in Theorem IV we replace (7) by the sum of two polynomials not of the same degree and neither of degree zero, we are led to the determination of a locus which is not generally bounded by circles. Compare 4 of the next to the last paragraph of this paper.

Any of these circles having no point in common with any other of these circles contains a number of roots of (11) equal to the multiplicity of the center of that circle as a root of (12).

A detailed proof of this theorem is quite simple. For any particular root z of (11) we consider the points $\{\alpha_i\}$ to coincide at $\bar{\alpha}$ on or in C . We have then

$$A_0(z - \bar{\alpha})^k + kA_1(z - \bar{\alpha})^{k-1} + \dots + k(k-1)\dots 2 \cdot 1A_k = 0.$$

Since $|\bar{\alpha} - \alpha| \leq r$, we know that z must lie in or on one of the circles determined. Conversely, any point z in or on one of these circles is a root of (11) for some choice of the $\{\alpha_i\}$; denote the center of the circle by β . We need merely choose $\alpha_1 = \alpha_2 = \dots = \alpha_k = \alpha + z - \beta$, which is a point on or within C .

If one of the circles C_1 has no point in common with any other of those circles, we may vary the α_i at will and no root of (11) can enter or leave C_1 ; these roots vary continuously if the α_i vary continuously. In particular if

$$\alpha_1 = \alpha_2 = \dots = \alpha_k = \alpha,$$

we have the number of roots of (11) proper to C_1 . It is similarly true that a number of these circles C_1, \dots, C_l having no point in common with any other of the circles contain a number of roots of (11) equal to the sum of the multiplicities of their centers considered as roots of (12).

Theorem VI enables us to give a result concerning the location of the roots of the sum of any number of arbitrary polynomials; $f(z)$ is to be chosen one of those polynomials of highest degree.

If in Theorem VI we place $A_0 = A_2 = A_3 = \dots = A_k = 0$, $A_1 \neq 0$, we have

A circle which contains all the roots of a polynomial contains also all the roots of its derivative.

This is essentially the theorem of Lucas:

*A convex polygon which contains all the roots of a polynomial contains also all the roots of its derivative.**

The similar results obtained from Theorem VI by placing $A_l \neq 0$ but all the other A_i zero ($l > 0$) give the theorem of Lucas for the other derivatives of $f(z)$.

Theorem I applies just as well to the integral of a polynomial as to its derivative, but for a result of the nature of those just given it is necessary to choose a *particular* integral. Thus we shall prove

THEOREM VII. *Let the common locus of all the roots of the polynomial*

$$f(z) = (z - \alpha_1)(z - \alpha_2)\dots(z - \alpha_k)$$

* In both of these statements as well as many other similar results, the term *contains* may be interpreted to include the possibility of points on the boundary or to exclude that possibility. Whichever interpretation is considered in the hypothesis, the conclusion will be true under that same interpretation.

be the interior and boundary of the circle C whose center is α and radius r . Then the locus of the roots of the polynomial

$$F(z) = \int_0^z f(z) dz$$

is composed of the interiors and boundaries of the $k + 1$ circles whose centers are $(1 - \omega)\alpha$ and radii $|1 - \omega|r$, where ω takes all the values of the $(k + 1)$ st roots of unity. Any of these circles which is entirely exterior to all the others contains precisely one root of $F(z)$.

If z is a root of $F(z)$ for particular values of the α_i , we must have for some choice of $\bar{\alpha} = \alpha_1 = \alpha_2 = \dots = \alpha_k$ in or on C ,

$$F(z) = \frac{(z - \bar{\alpha})^{k+1}}{k+1} - \frac{(-\bar{\alpha})^{k+1}}{k+1} = 0,$$

$$z = (1 - \omega)\bar{\alpha},$$

where $\omega^{k+1} = 1$. From this equation we have at once the desired inequality:

$$z - (1 - \omega)\alpha = (1 - \omega)(\bar{\alpha} - \alpha),$$

$$|z - (1 - \omega)\alpha| \leq |1 - \omega|r,$$

which shows that z is on or within one of the $k + 1$ circles. On the other hand, any point z on or within one of these $k + 1$ circles is a root of $F(z)$ for proper choice of the $\{\alpha_i\}$. In fact, if ω is the particular root of unity corresponding to that circle we have merely to place $\bar{\alpha} = \alpha_1 = \alpha_2 = \dots = \alpha_k$, where

$$\bar{\alpha} - \alpha = \frac{z - (1 - \omega)\alpha}{1 - \omega};$$

the exceptional case $\omega = 1$ is trivial.

The proof of the theorem is now complete except for the last sentence; this is proved in precisely the same manner as the corresponding statement in Theorem VI. Like Theorem VI, Theorem VII illustrates the remark made just previous to the statement of Theorem VI.

We shall next prove further results concerning the roots of the derivative of a polynomial, making continued use of the Lemma. The special case of the Lemma where m_1 and m_2 are positive is particularly simple; this special case leads to

THEOREM VIII. Denote by $g(z)$ the polynomial

$$(z - z_1)^{m_1} (z - z_2)^{m_2}$$

and by $m_1^{(n)} : m_2^{(n)}$ ($n = 1, 2, \dots, m$) the ratios in which the m distinct roots $z^{(n)}$ of $g^{(k)}(z)$ (the k th derivative of $g(z)$) divide the segment (z_1, z_2) .* Let the interiors and boundaries of circles C_1 and C_2 whose centers are α_1 and α_2 and radii r_1 and r_2 be the loci respectively of m_1 and m_2 roots of a polynomial $f(z)$ which

* The reader will easily prove from Rolle's Theorem that no point $z^{(n)}$ distinct from z_1 and z_2 can be a multiple root of $g^{(k)}(z)$.

Then we must have either

$$(z - \bar{\beta})^{q-1} = 0$$

or

$$ap(z - \bar{\beta})(z - \bar{\alpha})^{p-1} + q = 0.$$

If this latter equation is verified we find from Theorem III that there exists a point $\bar{\gamma}$ in or on C such that

$$\begin{aligned} ap(z - \bar{\gamma})^p + q &= 0, \\ z - \bar{\gamma} &= \omega \sqrt[p]{\frac{-q}{ap}}, \quad \text{where } \omega^p = 1. \end{aligned}$$

It therefore follows that

$$z - \left(\alpha + \omega \sqrt[p]{\frac{-q}{ap}} \right) = \bar{\gamma} - \alpha,$$

which is less than or equal to r in absolute value. Then all the roots of $f'(z)$ lie in C and the p circles of common radius r whose centers are the points

$$\alpha + \omega \sqrt[p]{\frac{-q}{ap}}.$$

We state this fact and its obvious converse in

THEOREM X. *Let the interior and boundary of the circle C whose center is α and radius r be the common locus of all the roots of the two polynomials*

$$\begin{aligned} P(z) &= a(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_p), & a &\neq 0, \\ Q(z) &= b(z - \beta_1)(z - \beta_2) \cdots (z - \beta_q), & b &\neq 0. \end{aligned}$$

Then the locus of the roots of the derivative of the function

$$e^{P(z)} \cdot Q(z)$$

is composed of C (unless $q = 1$) and the p circles whose common radius is r and whose centers are the points

$$\alpha + \omega \sqrt[p]{\frac{-q}{ap}},$$

where ω takes the values of the p th roots of unity. There are $q - 1$ roots of the derivative proper to C and one proper to each of the other p circles.

In the proof of this theorem we have used not the fact that the interior and boundary of C is the locus of the roots of $P(z)$ but that it is the locus of the roots of $P'(z)$, which gives in reality a more general result. Thus if $p = 1$, there is no restriction whatever on the root of $P(z)$.

We now prove a theorem which is a generalization of our result concerning the roots of (16) corresponding to our former generalization of Theorem VIII.

THEOREM XI. *Let the interior and boundary of a circle C whose center is α and radius is r be the common locus of the roots of the polynomials $\phi_1, \phi_2, \dots, \phi_q$ of respective degrees k_1, k_2, \dots, k_q . Then the locus of the roots of the polynomial*

$$(17) \quad \phi_1 \cdot \phi_2 \cdot \phi_3 \cdot \dots \cdot \phi_q - A\phi_1^{(n_1)} \cdot \phi_2^{(n_2)} \cdot \dots \cdot \phi_q^{(n_q)}, \quad n_i \leq k_i,$$

is composed of the circle C (unless $n_i = k_i$ for $i = 1, 2, \dots, q$) and the circles, $n_1 + n_2 + \dots + n_q$ in number, whose centers are the points

$$\alpha + \omega [k_1 \cdot (k_1 - 1) \dots (k_1 - n_1 + 1) \cdot k_2 \cdot (k_2 - 1) \dots (k_2 - n_2 + 1) \dots k_q \cdot (k_q - 1) \dots (k_q - n_q + 1) \cdot A]^{1/(n_1+n_2+\dots+n_q)}$$

where ω takes all possible values of an $(n_1 + n_2 + \dots + n_q)$ th root of unity, and whose common radius is r . There are

$$(k_1 + k_2 + \dots + k_q) - (n_1 + n_2 + \dots + n_q)$$

roots of (17) proper to C and one root proper to each of the $n_1 + n_2 + \dots + n_q$ other circles.

Under the hypothesis of the theorem, if z is a root of (17) there exist points $\alpha', \alpha'', \dots, \alpha^{(q)}$ on or within C and such that we have

$$(z - \alpha')^{k_1} (z - \alpha'')^{k_2} \dots (z - \alpha^{(q)})^{k_q} - [k_1 \cdot (k_1 - 1) \dots (k_1 - n_1 + 1) \cdot k_2 \cdot (k_2 - 1) \dots (k_2 - n_2 + 1) \cdot k_q \cdot (k_q - 1) \dots (k_q - n_q + 1) \cdot A \cdot (z - \alpha')^{k_1 - n_1} (z - \alpha'')^{k_2 - n_2} \dots (z - \alpha^{(q)})^{k_q - n_q}] = 0.$$

Then z must satisfy one of the equations

$$(z - \alpha')^{k_1 - n_1} (z - \alpha'')^{k_2 - n_2} \dots (z - \alpha^{(q)})^{k_q - n_q} = 0,$$

$$(z - \alpha')^{n_1} (z - \alpha'')^{n_2} \dots (z - \alpha^{(q)})^{n_q} - k_1 (k_1 - 1) \dots (k_q - n_q + 1) A = 0.$$

If z satisfies this latter equation we know from Theorem III that there exists a point $\bar{\alpha}$ on or within C and such that

$$(z - \bar{\alpha})^{n_1+n_2+\dots+n_q} = k_1 (k_1 - 1) \dots (k_q - n_q + 1) A.$$

That is, we have the equation

$$z - \{\alpha + \omega [k_1 (k_1 - 1) \dots (k_q - n_q + 1) A]^{1/(n_1+n_2+\dots+n_q)}\} = \bar{\alpha} - \alpha,$$

where $\omega^{n_1+n_2+\dots+n_q} = 1$. The absolute value of the right-hand member is not greater than r , so z must lie in or on one of the $n_1 + n_2 + \dots + n_q$ circles of the theorem.

The detailed proof of the remainder of Theorem XI now requires no further analytical work and is left to the reader.

The present paper aims merely to give some of the more immediate results that can be obtained from Theorem I. Further results can be obtained by extending in various directions:

1. Application of these methods to other and more general types of polynomials.

2. Extension of the results to other circular regions, namely, half-planes and the exteriors of circles.

3. Expression of the results in a manner which shall be independent of linear transformation of the complex variable.

4. Detailed determination and study of loci which naturally arise and which are of the same general nature as the locus determined in the lemma. For example, if the loci of α and β are two circular regions, determine the locus of the points z defined by

$$(z - \alpha)^m = A(z - \beta)^n.$$

For only the simplest polynomials $f(z, \alpha_i, \beta_i)$ do we get a number of circular regions as the locus of the roots of f when the loci of the roots of the α_i and β_i are circular regions. All the problems of the present paper have been characterized by a certain *linearity*, either the original problems themselves or those problems after a simple transformation. In more general problems, where the locus of the roots does not ordinarily consist of circular regions, the exact locus should be determined and also a simple even if rough approximation.

5. Application of these results and methods to the case of real polynomials.*

The present writer hopes to return later to a consideration of these questions.

* Thus the following theorem results from Theorem VI:

Let all the roots of a polynomial $f(z)$ of degree k lie in an interval of the axis of reals whose center is α and length l . Denote the roots of the polynomial

$$(1') \quad A_0 z^k + k A_1 z^{k-1} + \dots + k(k-1) \dots \cdot 2 \cdot 1 \cdot A_k$$

by z_1, z_2, \dots, z_k , which roots are supposed real. Then all the roots of the polynomial

$$(2') \quad A_0 f + A_1 f' + \dots + A_k f^{(k)}$$

lie in the intervals of the axis of reals whose centers are the points $\alpha + z_i$ ($i = 1, 2, \dots, k$) and whose common length is l . The number of roots of (2') proper to each of these intervals is the multiplicity of the corresponding z_i as a root of (1').

HARVARD UNIVERSITY,
CAMBRIDGE, MASS. (JUNE, 1921.)