KAKEYA'S PROBLEM ON THE ZEROS OF THE DERIVATIVE OF A POLYNOMIAL*

BY

MORRIS MARDEN†

1. **Introduction.** If all n zeros of a polynomial $f(z)$ of degree n lie in or on a circle K of radius R, then, according to the well known theorem of Gauss and Lucas,‡ all $n-1$ zeros of its derivative $f'(z)$ also lie in or on K. If only two zeros of $f(z)$ lie in or on K, then, according to a theorem stated by Alexander and proved by Kakeya and Szegö,§ at least one zero of $f'(z)$ lies in or on the concentric circle of radius $R \csc (\pi/n)$. If all but one of the zeros of $f(z)$ lie in or on K, then, according to a theorem due to Biernacki,|| at most one zero of $f'(z)$ lies outside of the concentric circle of radius $R(1+1/n)^{1/2}$. In general, according to a theorem stated by Kakeya,§ *if p zeros of a polynomial $f(z)$ of degree n, $(2 \leqq p \leqq n)$, lie in or on a circle of radius R, then at least $p-1$ zeros of its derivative lie in or on a concentric circle of radius $R\rho(n, p)$.*

The existence of a function $\rho(n, p)$ was proved by Kakeya§ in the general case. The actual computation of $\rho(n, p)$ seems, however, to have been made so far only in the three cases mentioned above; namely,

$$\rho(n, n) = 1, \qquad \rho(n, 2) \leqq \csc \pi/n, \qquad \rho(n, n-1) \leqq (1 + 1/n)^{1/2}.$$

Although in the present note the minimum value of $\rho(n, p)$ in the general case will not be determined, two inequalities for $\rho(n, p)$ will be established. First, for all n and p, $(2 \leqq p \leqq n)$,

$$\rho(n, p) \leqq \csc \frac{\pi}{2(n-p+1)}, \tag{1}$$

and, secondly, for at least p an even integer,¶

* Presented to the Society, September 1, 1936; received by the editors December 24, 1937 and, in revised form, May 19, 1938.

† The results of this paper were obtained under a grant for summer research from the Wisconsin Alumni Research Foundation.

‡ See references in M. Marden, American Mathematical Monthly, vol. 42 (1935), pp. 278–279.

§ J. W. Alexander, Annals of Mathematics, (2), vol. 17 (1915), p. 18; S. Kakeya, Tôhoku Mathematical Journal, vol. 11 (1917), pp. 5–16, especially p. 9; G. Szegö, Mathematische Zeitschrift, vol. 13 (1932), pp. 28–55.

|| M. Biernacki, Bulletin de l'Académie Polonaise, 1927, pp. 660–675; See also J. L. Walsh, these Transactions, vol. 24 (1922), p. 37.

¶ It is to be noted that for $p=n$, $\csc \pi/[2(n-p+1)]=1=(2-p/n)^{1/2}$; and, for $p<n$, $\csc \pi/[2(n-p+1)] \geqq 2^{1/2} > (2-p/n)^{1/2}$.

$$\rho(n, p) \geqq (2 - p/n)^{1/2}. \tag{2}$$

The second inequality may be proved simply by exhibiting a polynomial of degree n which has $p=2m$ zeros in or on the unit circle and of which the derivative has at least $p-1$ zeros in or on the circle $|z| = (2-p/n)^{1/2}$. Such a polynomial is

$$f(z) = \left[z^2 - 2z\left(\frac{n}{2n-p}\right)^{1/2} + 1\right]^{p/2}\left[z - \frac{1}{p}(n(2n-p))^{1/2}\right]^{n-p};$$

for, it has zeros of multiplicity $p/2$ on the unit circle at the points

$$z = \left(\frac{n}{2n-p}\right)^{1/2} \pm i\left(\frac{n-p}{2n-p}\right)^{1/2},$$

and its derivative has zeros of multiplicity $(p-2)/2$ at these points and a double zero at the point $z=(2-p/n)^{1/2}$.

The proof of the first inequality, however, will require the establishment of an identity (apparently new) relating any p zeros of a polynomial

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

with any $(n-p+1)$ zeros of its derivative which are distinct from the p given zeros of $f(z)$. The identity is a generalization of the well known formula

$$\sum_{j=1}^{n} \frac{1}{\beta - \alpha_j} = 0$$

relating the n zeros of $f(z)$ with any one zero β of $f'(z)$ which is not a zero of $f(z)$.

The identity in question is derived in §2 and applied to the proof of inequality (1) in §3. In §4, the relation of this inequality to one given by Fekete is discussed. Finally, in §5, the inequality is used to obtain a sufficient condition for a polynomial to be at most p-valent in a given circle or other convex region.

2. **An identity.** The identity mentioned above is described in the following theorem:

THEOREM 1. *If the* $n+1$ *complex numbers*

$$\alpha_1, \alpha_2, \cdots, \alpha_p; \beta_1, \beta_2, \cdots, \beta_q, \qquad 2 \leqq p \leqq n,\ q = n - p + 1,$$

are distinct and if all of the α_j *are zeros of a polynomial* $f(z)$ *of degree* n *and all of the* β_k *are zeros of its derivative* $f'(z)$, *then*

$$\sum \frac{D_{j_1 j_2 \cdots j_q}}{(\beta_1 - \alpha_{j_1})(\beta_2 - \alpha_{j_2}) \cdots (\beta_q - \alpha_{j_q})} = 0, \tag{3}$$

where $j_1, j_2, \cdots, j_q$ *run independently from* 1 *to* p, *where*

$$D_{j_1 j_2 \cdots j_q} = \prod_{m=1}^{p} (\delta_{mj_1} + \delta_{mj_2} + \cdots + \delta_{mj_q})!,$$

and where $\delta_{mj} = 1$ *or* 0 *according as* $j = m$ *or* $j \neq m$.

To prove Theorem 1, we shall let

$$P(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_p).$$

Then there exist q constants $a_0, a_1, \cdots, a_{q-1}$, not all zero, such that

$$f(z) = (a_0 + a_1 z + \cdots + a_{q-1} z^{q-1}) P(z).$$

These constants satisfy the system of q homogeneous linear equations

$$f'(\beta_j) = a_0 \frac{d}{d\beta_j} P(\beta_j) + a_1 \frac{d}{d\beta_j} [\beta_j P(\beta_j)] + \cdots + a_{q-1} \frac{d}{d\beta_j} [\beta_j^{q-1} P(\beta_j)] = 0,$$
$$j = 1, 2, \cdots, q,$$

of which system the determinant

$$\Delta(\beta_1, \beta_2, \cdots, \beta_q) = \begin{vmatrix} \frac{d}{d\beta_1} P(\beta_1) & \frac{d}{d\beta_1} [\beta_1 P(\beta_1)] & \cdots & \frac{d}{d\beta_1} [\beta_1^{q-1} P(\beta_1)] \\ \frac{d}{d\beta_2} P(\beta_2) & \frac{d}{d\beta_2} [\beta_2 P(\beta_2)] & \cdots & \frac{d}{d\beta_2} [\beta_2^{q-1} P(\beta_2)] \\ \cdot & \cdot & \cdots & \cdot \\ \frac{d}{d\beta_q} P(\beta_q) & \frac{d}{d\beta_q} [\beta_q P(\beta_q)] & \cdots & \frac{d}{d\beta_q} [\beta_q^{q-1} P(\beta_q)] \end{vmatrix} \tag{4.1}$$

must therefore vanish.

Defining $V(\beta_1, \beta_2, \cdots, \beta_q)$ as the Vandermondian determinant

$$V(\beta_1, \beta_2, \cdots, \beta_q) = \begin{vmatrix} 1 & \beta_1 & \beta_1^2 & \cdots & \beta_1^{q-1} \\ 1 & \beta_2 & \beta_2^2 & \cdots & \beta_2^{q-1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 1 & \beta_q & \beta_q^2 & \cdots & \beta_q^{q-1} \end{vmatrix} = \prod_{j=1}^{q} \prod_{k=j+1}^{q} (\beta_k - \beta_j), \tag{4.2}$$

we may write

$$\Delta(\beta_1, \beta_2, \cdots, \beta_q) = \frac{\partial^q}{\partial\beta_1 \partial\beta_2 \cdots \partial\beta_q} [P(\beta_1) P(\beta_2) \cdots P(\beta_q) V(\beta_1, \beta_2, \cdots, \beta_q)],$$

and hence

$$\frac{\partial^{k_1+k_2+\cdots+k_q-q}\Delta}{\partial\beta_1^{k_1}\partial\beta_2^{k_2}\cdots\partial\beta_q^{k_q}} = \frac{\partial^{k_1+k_2+\cdots+k_q}}{\partial\beta_1^{k_1}\partial\beta_2^{k_2}\cdots\partial\beta_q^{k_q}}\left[P(\beta_1)P(\beta_2)\cdots P(\beta_q)V\right].$$

The right-hand side of this equation may be evaluated by Leibniz' rule for differentiating a product, as follows. First,

$$\frac{\partial^{k_1}}{\partial\beta_1^{k_1}}\left[P(\beta_1)V\right] = \sum_{j_1=0}^{k_1} C_{k_1,j_1}P^{(k_1-j_1)}(\beta_1)\frac{\partial^{j_1}V}{\partial\beta_1^{j_1}},$$

where $C_{k_1,j_1}=k_1!/j_1!(k_1-j_1)!$ and $C_{k_1,0}=1$. If, now, we assume that for some fixed value of m, $(1 \leqq m \leqq q)$,

$$\frac{\partial^{k_1+k_2+\cdots+k_m}}{\partial\beta_1^{k_1}\partial\beta_2^{k_2}\cdots\partial\beta_m^{k_m}}\left[P(\beta_1)P(\beta_2)\cdots P(\beta_m)V\right]$$
$$= \sum_{j_1=0}^{k_1}\sum_{j_2=0}^{k_2}\cdots\sum_{j_m=0}^{k_m}\left\{\prod_{i=1}^{m} C_{k_i,j_i}P^{(k_i-j_i)}(\beta_i)\frac{\partial^{j_i}}{\partial\beta_i^{j_i}}\right\}V,$$

then

$$\frac{\partial^{k_1+k_2+\cdots+k_{m+1}}}{\partial\beta_1^{k_1}\partial\beta_2^{k_2}\cdots\partial\beta_{m+1}^{k_{m+1}}}\left[P(\beta_1)P(\beta_2)\cdots P(\beta_{m+1})V\right]$$
$$= \sum_{j_1=0}^{k_1}\cdots\sum_{j_m=0}^{k_m}\left\{\prod_{i=1}^{m} C_{k_i,j_i}P^{(k_i-j_i)}(\beta_i)\frac{\partial^{j_i}}{\partial\beta_i^{j_i}}\right\}\frac{\partial^{k_{m+1}}}{\partial\beta_{m+1}^{k_{m+1}}}P(\beta_{m+1})V$$
$$= \sum_{j_1=0}^{k_1}\cdots\sum_{j_m=0}^{k_m}\left\{\prod_{i=1}^{m} C_{k_i,j_i}P^{(k_i-j_i)}(\beta_i)\frac{\partial^{j_i}}{\partial\beta_i^{j_i}}\right\}$$
$$\cdot\sum_{j_{m+1}=0}^{k_{m+1}} C_{k_{m+1},j_{m+1}}P^{(k_{m+1}-j_{m+1})}(\beta_{m+1})\frac{\partial^{j_{m+1}}V}{\partial\beta_{m+1}^{j_{m+1}}}$$
$$= \sum_{j_1=0}^{k_1}\sum_{j_2=0}^{k_2}\cdots\sum_{j_{m+1}=0}^{k_{m+1}}\left\{\prod_{i=1}^{m+1} C_{k_i,j_i}P^{(k_i-j_i)}(\beta_i)\frac{\partial^{j_i}}{\partial\beta_i^{j_i}}\right\}V.$$

We may conclude by mathematical induction, therefore, that

$$\frac{\partial^{k_1+k_2+\cdots+k_q-q}\Delta}{\partial\beta_1^{k_1-1}\partial\beta_2^{k_2-1}\cdots\partial\beta_q^{k_q-1}} \tag{5.1}$$
$$= \sum_{j_1=0}^{k_1}\sum_{j_2=0}^{k_2}\cdots\sum_{j_q=0}^{k_q}\left\{\prod_{i=1}^{q} C_{k_i,j_i}P^{(k_i-j_i)}(\beta_i)\frac{\partial^{j_i}}{\partial\beta_i^{j_i}}\right\}V.$$

It is furthermore clear that $\Delta(\beta_1, \beta_2, \cdots, \beta_q)$ is a polynomial of degree $n-1$ in each β_j. Since Δ vanishes when any two β_j are equated, Δ must have V as a factor. Hence the quotient

$$\Phi(\beta_1, \beta_2, \cdots, \beta_q) = \frac{\Delta(\beta_1, \beta_2, \cdots, \beta_q)}{V(\beta_1, \beta_2, \cdots, \beta_q)}$$

is a polynomial of degree

$$(n-1)-(q-1)=n-q=p-1$$

in each β_j, and, as is evident from formulas (4.1) and (4.2), it is symmetric in the β_j.

Since $P(z)$ is a polynomial of degree p and has no multiple zeros, it is true, according to Lagrange's interpolation formula that

$$\frac{\Phi(\beta_1, \beta_2, \cdots, \beta_q)}{P(\beta_1)} = \sum_{j_1=1}^{p} \frac{\Phi(\alpha_{j_1}, \beta_2, \beta_3, \cdots, \beta_q)}{P'(\alpha_{j_1})(\beta_1 - \alpha_{j_1})}.$$

If, now, it be assumed that, for m any fixed positive integer less than q,

$$\frac{\Phi(\beta_1, \beta_2, \cdots, \beta_q)}{P(\beta_1)P(\beta_2)\cdots P(\beta_m)} = \sum_{j_1=1}^{p}\sum_{j_2=1}^{p}\cdots\sum_{j_m=1}^{p} \frac{\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_m}, \beta_{m+1}, \beta_{m+2}, \cdots, \beta_q)}{P'(\alpha_{j_1})P'(\alpha_{j_2})\cdots P'(\alpha_{j_m})(\beta_1-\alpha_{j_1})(\beta_2-\alpha_{j_2})\cdots(\beta_m-\alpha_{j_m})},$$

then again by Lagrange's formula

$$\frac{\Phi(\beta_1, \beta_2, \cdots, \beta_q)}{P(\beta_1)P(\beta_2)\cdots P(\beta_{m+1})} = \sum_{j_1=1}^{p}\sum_{j_2=1}^{p}\cdots\sum_{j_m=1}^{p} \frac{1}{P'(\alpha_{j_1})P'(\alpha_{j_2})\cdots P'(\alpha_{j_m})(\beta_1-\alpha_{j_1})(\beta_2-\alpha_{j_2})\cdots(\beta_m-\alpha_{j_m})} \cdot \sum_{j_{m+1}=1}^{p} \frac{\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_{m+1}}, \beta_{m+2}, \cdots, \beta_q)}{P'(\alpha_{j_{m+1}})(\beta_{m+1} - \alpha_{j_{m+1}})}.$$

It follows then by mathematical induction that

$$\frac{\Phi(\beta_1, \beta_2, \cdots, \beta_q)}{P(\beta_1)P(\beta_2)\cdots P(\beta_q)} = \sum_{j_1=1}^{p}\sum_{j_2=1}^{p}\cdots\sum_{j_q=1}^{p} \frac{\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q})}{P'(\alpha_{j_1})\cdots P'(\alpha_{j_q})(\beta_1-\alpha_{j_1})\cdots(\beta_q-\alpha_{j_q})}. \tag{5.2}$$

Let us next compute the value of $\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q})$ for a given set A of the α_{j_i}.

First, let us consider the case that no two α_{j_i} of set A are equal. Then from formula (4.2) it follows that

$$V(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) \neq 0;$$

and, since the α_{j_i} are zeros of $P(z)$, it follows from (4.1) or (5.1) that

$$\Delta(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = P'(\alpha_{j_1})P'(\alpha_{j_2})\cdots P'(\alpha_{j_q})V(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) \neq 0$$

and hence that

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = P'(\alpha_{j_1})P'(\alpha_{j_2}) \cdots P'(\alpha_{j_q}).$$

Secondly, let us consider the case that in the set A, $\alpha_{j_1}=\alpha_{j_2}= \cdots =\alpha_{j_\mu}$, $(\mu \leqq q)$, but $\alpha_{j_\mu}, \alpha_{j_{\mu+1}}, \cdots, \alpha_{j_q}$ are distinct. From formulas (4.1) and (4.2) it then follows that the derivatives

$$\left[\frac{\partial^{k_1}}{\partial\beta_1^{k_1}} \frac{\partial^{k_2}}{\partial\beta_2^{k_2}} \cdots \frac{\partial^{k_\mu}}{\partial\beta_\mu^{k_\mu}} V\right]_A, \tag{6.1}$$

$$\left[\frac{\partial^{k_1}}{\partial\beta_1^{k_1}} \frac{\partial^{k_2}}{\partial\beta_2^{k_2}} \cdots \frac{\partial^{k_\mu}}{\partial\beta_\mu^{k_\mu}} \Delta\right]_A \tag{6.2}$$

vanish whenever two or more k_i are equal and, therefore, whenever

$$k_1 + k_2 + \cdots + k_\mu \leqq 0 + 1 + 2 + \cdots + (\mu - 1) = \mu(\mu - 1)/2$$

unless $(k_1, k_2, \cdots, k_\mu)$ is the set K_1: $(0, 1, \cdots, \mu-1)$ or a set obtainable by merely permuting the numbers of the set K_1. These $\mu!$ sets will be referred to hereafter as the sets K.

In the neighborhood of the point A,

$$\Phi(\beta_1, \beta_2, \cdots, \beta_q) = \frac{\sum \zeta_1^{k_1}\zeta_2^{k_2} \cdots \zeta_\mu^{k_\mu}\left\{\left[\frac{\partial^{k_1}}{\partial\beta_1^{k_1}} \frac{\partial^{k_2}}{\partial\beta_2^{k_2}} \cdots \frac{\partial^{k_\mu}}{\partial\beta_\mu^{k_\mu}} \Delta\right]_A + \epsilon_{k_1k_2\cdots k_q}\right\}}{\sum \zeta_1^{k_1}\zeta_2^{k_2} \cdots \zeta_\mu^{k_\mu}\left\{\left[\frac{\partial^{k_1}}{\partial\beta_1^{k_1}} \frac{\partial^{k_2}}{\partial\beta_2^{k_2}} \cdots \frac{\partial^{k_\mu}}{\partial\beta_\mu^{k_\mu}} V\right]_A + \epsilon'_{k_1k_2\cdots k_q}\right\}}$$

where $\zeta_j=\beta_j-\alpha_{j_i}$,

$$[\epsilon_{k_1k_2\cdots k_q}]_A = [\epsilon'_{k_1k_2\cdots k_q}]_A = 0,$$

and both sums are taken over all sets K. Furthermore, since changing from one set K to another set K merely multiples both derivatives (6.1) and (6.2) by one or both by minus one, we may write

$$\Phi(\beta_1, \beta_2, \cdots, \beta_q) = \frac{\sum \zeta_1^{k_1}\zeta_2^{k_2} \cdots \zeta_\mu^{k_\mu}\left\{\pm\left[\frac{\partial}{\partial\beta_2} \frac{\partial^2}{\partial\beta_3^2} \cdots \frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}} \Delta\right]_A + \epsilon_{k_1k_2\cdots k_q}\right\}}{\sum \zeta_1^{k_1}\zeta_2^{k_2} \cdots \zeta_\mu^{k_\mu}\left\{\pm\left[\frac{\partial}{\partial\beta_2} \frac{\partial^2}{\partial\beta_3^2} \cdots \frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}} V\right]_A + \epsilon'_{k_1k_2\cdots k_q}\right\}}.$$

If therefore for a given path of approach of the point $(\beta_1, \beta_2, \cdots, \beta_q)$ to A $\lim (\zeta_i/\zeta_1)=\eta_i$, then

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = \frac{\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}\Delta\right]_A \sum (\pm)\eta_2^{k_2}\eta_3^{k_3}\cdots\eta_\mu^{k_\mu}}{\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}V\right]_A \sum (\pm)\eta_2^{k_2}\eta_3^{k_3}\cdots\eta_\mu^{k_\mu}}$$

$$= \frac{\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}\Delta\right]_A}{\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}V\right]_A}.$$

Finally, according to formula (5.1),

$$\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}\Delta\right]_A = \sum_{j_1=0}^{1}\sum_{j_2=0}^{2}\cdots\sum_{j_\mu=0}^{\mu}\sum_{j_\mu+1=0}^{1}\cdots\sum_{j_q=0}^{1} C_{2,j_2}C_{3,j_3}\cdots C_{\mu,j_\mu}$$
$$\cdot P^{(1-j_1)}(\alpha_{j_1})P^{(2-j_2)}(\alpha_{j_2})\cdots P^{(\mu-j_\mu)}(\alpha_{j_\mu})P^{(1-j_{\mu+1})}(\alpha_{j_{\mu+1}})$$
$$\cdots P^{(1-j_q)}(\alpha_{j_q})\left[\frac{\partial^{j_1}}{\partial\beta_1^{j_1}}\frac{\partial^{j_2}}{\partial\beta_2^{j_2}}\cdots\frac{\partial^{j_q}}{\partial\beta_q^{j_q}}V\right]_A.$$

Since the α_{j_i} are zeros of $P(z)$,

$$\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}\Delta\right]_A = \sum_{j_2=0}^{1}\sum_{j_3=0}^{2}\cdots\sum_{j_\mu=0}^{\mu-1} C_{2,j_2}C_{3,j_3}\cdots C_{\mu,j_\mu}$$
$$\cdot P'(\alpha_{j_1})P^{(2-j_2)}(\alpha_{j_2})\cdots P^{(\mu-j_\mu)}(\alpha_{j_\mu})P'(\alpha_{j_{\mu+1}})$$
$$\cdots P'(\alpha_{j_q})\left[\frac{\partial^{j_2}}{\partial\beta_2^{j_2}}\frac{\partial^{j_3}}{\partial\beta_3^{j_3}}\cdots\frac{\partial^{j_\mu}}{\partial\beta_\mu^{j_\mu}}V\right]_A.$$

By use of our above remarks on the vanishing of the derivative (6.1), this expression reduces further to

$$\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}\Delta\right]_A = C_{2,1}C_{3,2}\cdots C_{\mu,\mu-1}P'(\alpha_{j_1})P'(\alpha_{j_2})$$
$$\cdots P'(\alpha_{j_q})\left[\frac{\partial}{\partial\beta_2}\frac{\partial^2}{\partial\beta_3^2}\cdots\frac{\partial^{\mu-1}}{\partial\beta_\mu^{\mu-1}}V\right]_A;$$

and, consequently,

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = \mu!P'(\alpha_{j_1})P'(\alpha_{j_2})\cdots P'(\alpha_{j_q}).$$

Thirdly, let us consider the case that all of the α_{j_i} in the set A are distinct except for μ of the α_{j_i} which are equal to one another, these μ of the α_{j_i} not being necessarily the first μ of the α_{j_i}. Then, due to the symmetry of the function $\Phi(\beta_1, \beta_2, \cdots, \beta_q)$ in the β_j, we see that the result obtained in the second case holds here also; namely,

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = \mu! P'(\alpha_{j_1}) P'(\alpha_{j_2}) \cdots P'(\alpha_{j_q}).$$

Lastly, let us consider the most general case. Because of the symmetry of $\Phi(\beta_1, \beta_2, \cdots, \beta_q)$ in the β_j, we may, without loss of generality, assume that $1 \leqq j_1 \leqq j_2 \leqq \cdots \leqq j_q \leqq p$. Let us suppose concerning the set A that

$$\begin{aligned}
\alpha_{j_1} &= \alpha_{j_2} = \cdots = \alpha_{j_{\sigma_1}} = A_1, \\
\alpha_{j_{\sigma_1+1}} &= \alpha_{j_{\sigma_1+2}} = \cdots = \alpha_{j_{\sigma_2}} = A_2, \\
&\cdots\cdots\cdots\cdots\cdots, \\
\alpha_{j_{\sigma_{s-1}+1}} &= \alpha_{j_{\sigma_{s-1}+2}} = \cdots = \alpha_{j_{\sigma_s}} = A_s
\end{aligned}$$

where $0 = \sigma_0 < \sigma_1 < \cdots < \sigma_s = q$ and all the A_i are distinct. Let $\mu_i = \sigma_i - \sigma_{i-1}$. Then, exactly as in the second case, one concludes that

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = \frac{\left[\left(\prod_{m=1}^{s} \frac{\partial}{\partial \beta_{\sigma_{m-1}+2}} \frac{\partial^2}{\partial \beta^2_{\sigma_{m-1}+3}} \cdots \frac{\partial^{\mu_m - 1}}{\partial \beta_{\sigma_m}^{\mu_m - 1}}\right) \Delta\right]_A}{\left[\left(\prod_{m=1}^{s} \frac{\partial}{\partial \beta_{\sigma_{m-1}+2}} \frac{\partial^2}{\partial \beta^2_{\sigma_{m-1}+3}} \cdots \frac{\partial^{\mu_m - 1}}{\partial \beta_{\sigma_m}^{\mu_m - 1}}\right) V\right]_A},$$

that

$$\begin{aligned}
&\left[\left(\prod_{m=1}^{s} \frac{\partial}{\partial \beta_{\sigma_{m-1}+2}} \frac{\partial^2}{\partial \beta^2_{\sigma_{m-1}+3}} \cdots \frac{\partial^{\mu_m - 1}}{\partial \beta_{\sigma_m}^{\mu_m - 1}}\right) \Delta\right]_A \\
&\qquad = \prod_{m=1}^{s} (\mu_m!) P'(\alpha_{j_1}) P'(\alpha_{j_2}) \\
&\qquad\qquad \cdots P'(\alpha_{j_q}) \left[\left(\prod_{m=1}^{s} \frac{\partial}{\partial \beta_{\sigma_{m-1}+2}} \frac{\partial^2}{\partial \beta^2_{\sigma_{m-1}+3}} \cdots \frac{\partial^{\mu_m - 1}}{\partial \beta_{\sigma_m}^{\mu_m - 1}}\right) V\right]_A,
\end{aligned}$$

and therefore that

$$\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q}) = \prod_{m=1}^{s} (\mu_m!) P'(\alpha_{j_1}) P'(\alpha_{j_2}) \cdots P'(\alpha_{j_q}).$$

The substitution into equation (5.2) of the above value of $\Phi(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_q})$ and the use of the relation $\mu_m = \sum_{i=1}^{q} \delta_{m j_i}$ complete the proof of Theorem 1.

In the foregoing proof of Theorem 1, the $n+1$ numbers $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_q$ were assumed to be all distinct. Since the left-hand side of equation (3) is a function which is continuous in all the α_j and β_k except for the values $\alpha_j = \beta_k$, $(j = 1, 2, \cdots, p;\ k = 1, 2, \cdots, q)$, an identity similar to (3) may be derived as a limiting case of (3), in the event that not all of the α_j are distinct and that not all of the β_k are distinct, provided that no α_j is a β_k.

Suppose, for example, that

$$\alpha_1 = \alpha_2 = \cdots = \alpha_t, \qquad t \leqq p,$$

but that $\alpha_t, \alpha_{t+1}, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_q$ are distinct. Then equation (3) may be written as

$$\sum_{j_1=t}^{p} \sum_{j_2=t}^{p} \cdots \sum_{j_q=t}^{p} \frac{E_{j_1 j_2 \cdots j_q}}{(\beta_1 - \alpha_{j_1})(\beta_2 - \alpha_{j_2}) \cdots (\beta_q - \alpha_{j_q})} = 0 \tag{7}$$

where $E_{j_1 j_2 \cdots j_q}$ is a constant which will now be determined. If α_t occurs exactly λ times in the denominator of a given term of equation (7), for example, in the product

$$(\beta_1 - \alpha_t)(\beta_2 - \alpha_t) \cdots (\beta_\lambda - \alpha_t), \tag{8}$$

that term may be considered as the limit of the sum of all terms of equation (3) in the denominators of which occur the products

$$(\beta_1 - \alpha_{j_1})(\beta_2 - \alpha_{j_2}) \cdots (\beta_\lambda - \alpha_{j_\lambda})$$

where the α_{j_i} are selected in all possible ways from the set

$$(\alpha_1, \alpha_2, \cdots, \alpha_t).$$

Suppose κ_1 α_1's, κ_2 α_2's, $\cdots$, and κ_t α_t's, where $\kappa_j \geqq 0$, all j, and

$$\kappa_1 + \kappa_2 + \cdots + \kappa_t = \lambda \tag{9}$$

are selected. There are in (3)

$$\frac{\lambda!}{\kappa_1!\kappa_2! \cdots \kappa_t!}$$

terms which contain the chosen α_i and, according to Theorem 1, each of these terms will have as a factor of the numerator coefficient D the product

$$\kappa_1!\kappa_2! \cdots \kappa_t!.$$

Hence, the factor $\lambda!$ occurs in the numerator of the limit of the sum of such terms. The set of nonnegative integers $(\kappa_1, \kappa_2, \cdots, \kappa_t)$ may, in addition, be selected subject to the condition (9) in $C_{t+\lambda-1,\lambda}$ ways. Hence, the factor corresponding to (8) in the numerator of the given term of (7) will be

$$\lambda! C_{t+\lambda-1,\lambda} = t(t+1)(t+2) \cdots (t+\lambda-1).$$

On the other hand, suppose that $\beta_1 = \beta_2 = \cdots = \beta_u$ but that $\beta_u, \beta_{u+1}, \cdots, \beta_q$ are distinct. Then, since the number of terms of (3) in which $\beta_1, \beta_2, \cdots, \beta_u$ are associated with δ_1 α_1's, δ_2 α_2's, $\cdots$, δ_p α_p's, where $\delta_1 + \delta_2 + \cdots + \delta_p = u$, is

$$\frac{u!}{\delta_1!\delta_2! \cdots \delta_p!},$$

that number of terms coalesce to form the single corresponding term of the limit of (3).

Thus the following corollary is evident:

COROLLARY. *Among the* $r+s$ *distinct numbers*

$$A_1, A_2, \cdots, A_r, \qquad B_1, B_2, \cdots, B_s$$

let each A_j *be a zero, of multiplicity at least* p_j, *of a polynomial* $f(z)$ *of degree* n, *and each* B_k *a zero, of multiplicity at least* q_k, *of the derivative of* $f(z)$ *where*

$$2 \leqq p_1 + p_2 + \cdots + p_r = p \leqq n$$

and

$$q_1 + q_2 + \cdots + q_s = q = n - p + 1.$$

Then the A_j *and* B_k *satisfy the relation*

$$\sum \prod_{j=1}^{s} \prod_{k=1}^{r} \frac{\mu_k ! q_j ! C_{p_k+\mu_k-1,\mu_k}}{\nu_{jk}!(B_j - A_k)^{\nu_{jk}}} = 0$$

where the sum is formed for all ν_{jk}, $(j=1, 2, \cdots, s;\ k=1, 2, \cdots, r)$, *such that* $\nu_{jk}=0, 1, 2, \cdots, q_j$ *and* $\nu_{j1}+\nu_{j2}+\cdots+\nu_{jr}=q_j$, *and where* $\mu_k=\nu_{1k}+\nu_{2k}+\cdots+\nu_{sk}$.

3. **Proof of inequality** (2). Theorem 1 and its corollary will now be applied to the establishing of the following theorem:

THEOREM 2. *If a polynomial* $f(z)$ *of degree* n, $(n\geqq 2)$, *has* p, $(p\geqq 2)$, *zeros in or on a circle* K *of radius* R, *then its derivative* $f'(z)$ *has at least* $p-1$ *zeros in or on the concentric circle* K' *of radius*

$$R' = R \csc \frac{\pi}{2(n - p + 1)} \cdot$$

For the proof of Theorem 2, it may be assumed without loss of generality that K is the unit circle $|z| = 1$.

Let $\alpha_1, \alpha_2, \cdots, \alpha_p$ be the p given zeros of $f(z)$, and let $\beta_1, \beta_2, \cdots, \beta_{n-1}$ be all $n-1$ zeros of $f'(z)$, the subscripts on the β_j being chosen so that

$$|\beta_1| \geqq |\beta_2| \geqq \cdots \geqq |\beta_{n-1}|.$$

If $|\beta_q| \leqq 1$, where $q=n-p+1$, then also $|\beta_j| \leqq 1$, $(j=q+1, q+2, \cdots, n-1)$; that is to say, at least $(n-q)=(p-1)$ of the β_j will lie in the unit circle and therefore in the circle K'.

If $|\beta_q| >1$, then likewise $|\beta_j| >1$, $(j=1, 2, \cdots, q-1)$. As $|\alpha_k| \leqq 1$ for all k, no β_j, $(j=1, 2, \cdots, q)$, will be an α_k; hence either Theorem 1 or its corollary may be used. Let ϕ_j be the angle subtended by the circle K in the point β_j, and let α_j' denote the α_k corresponding to a given β_j such that

$$0 \leqq \arg \frac{\beta_j - \alpha_j'}{\beta_j - \alpha_k} \leqq \phi_j \leqq \phi_q < \pi$$

for all $j=1, 2, \cdots, q$ and all $k=1, 2, \cdots, p$. It follows that

$$0 \leqq \arg \frac{\prod_{j=1}^{q} (\beta_j - \alpha_j')}{\prod_{j=1}^{q} (\beta_j - \alpha_{k_j})} \leqq (q - \delta)\phi_q$$

where δ, $(0 \leqq \delta \leqq q)$, denotes the number of factors common to the two products

$$\prod_{j=1}^{q} (\beta_j - \alpha_j'), \qquad \prod_{j=1}^{q} (\beta_j - \alpha_{k_j}).$$

If, therefore, $\phi_q < \pi/q$, each term in the sum obtained on multiplying the left-hand side of (3) by

$$\prod_{j=1}^{q} (\beta_j - \alpha_j')$$

could be represented by a vector drawn from the origin to a point lying in the angular opening

$$0 \leqq \arg z < \pi;$$

hence the left-hand side of (3) would not vanish. As this result would contradict Theorem 1, it follows that $\phi_q \geqq \pi/q$; that is to say, the $p-1$ zeros of $f'(z)$ $\beta_q, \beta_{q+1}, \cdots, \beta_{n-1}$ lie in or on a circle K' concentric with K and of radius

$$R' = \csc \frac{\pi}{2q} = \csc \frac{\pi}{2(n - p + 1)}.$$

The above method of proof may also be used, with little change, in the case that K is a convex region not necessarily a circle. The corresponding result may be stated as follows:

THEOREM 2'. *If a polynomial $f(z)$ of degree n, $(n \geqq 2)$, has p, $(p \geqq 2)$, zeros in a convex region K, its derivative has at least $p-1$ zeros in the star-shaped region K' consisting of all points of the plane from which K subtends an angle of not less than $\pi/(n-p+1)$ radians.*

Theorem 2 or Theorem 2' does not furnish, however, the least number

$\rho(n, p)$ as defined in §1. This is clear from the fact that, in general, the quantity δ used in the proof takes on values in addition to 0 and that, therefore ϕ_q must be actually greater than π/q in order for the left-hand side of (3) to vanish.

The same is clear from the facts that, although for $p=n$

$$\csc \frac{\pi}{2(n-p+1)} = 1 = \rho(n, n),$$

nevertheless for $p=2$ and $n \geqq 3$

$$\csc \frac{\pi}{2(n-p+1)} > \csc \pi/n \geqq \rho(n, 2),$$

and for* $p=n-1$ and $n \geqq 2$

$$\csc \frac{\pi}{2(n-p+1)} = 2^{1/2} > (1 + 1/n)^{1/2} \geqq \rho(n, n-1).$$

4. **Relation to a theorem of Fekete.** The inequality

$$\rho(n, 2) \leqq \csc \pi/n$$

was proved by Szegö as a consequence of the following theorem of Grace and Heawood:† *If a and b are two distinct zeros of a polynomial $f(z)$ of degree n, at least one zero of the derivative of $f(z)$ lies in or on the circle*

$$\left| z - \frac{a+b}{2} \right| = \left| \frac{a-b}{2} \right| \cot \pi/n. \tag{10}$$

Szegö showed that if the a and b are allowed to vary independently in and on the unit circle, the envelope of circle (10) is the circle $|z| = \csc \pi/n$.

A similar relation will now be proved to hold between Theorem 2 and the following theorem:

THEOREM 3. *If a and b are respectively k-fold and l-fold zeros of a polynomial $f(z)$ of degree n, then at least one zero (different from a and b) of the derivative lies in or on the circle*

$$\left| z - \frac{a+b}{2} \right| = \left| \frac{a-b}{2} \right| \cot \frac{\pi}{2(n+1-k-l)}. \tag{11}$$

---

* For $p=n-1$ and $n \geqq 5$, $\csc \pi/[2(n-p+1)] = 2^{1/2} > 1+2/n$, where $1+2/n$ is a limit obtainable from a theorem due to Walsh. See J. L. Walsh, these Transactions, vol. 24 (1922), p. 37, and also Biernacki, Bulletin de l'Académie Polonaise, 1927, p. 121.

† J. H. Grace, Proceedings of the Cambridge Philosophical Society, vol. 11 (1901), pp. 352–357; P. J. Heawood, Quarterly Journal of Mathematics, vol. 38 (1907), pp. 84–107.

This theorem, a generalization of one due to Fekete,* is an immediate result of the lemma:†

If $P(z)$ is a polynomial of degree $\nu \geqq 1$, if $\phi(z)$ is a function real, continuous, nonnegative, and not identically vanishing on the interval $(-1, 1)$ *of the real axis, and if*

$$\int_{-1}^{1} \phi(z)P(z)dz = 0,$$

then $P(z)$ vanishes in at least one point in which the segment $(-1, 1)$ *subtends an angle of not less than π/ν.*

In the proof of Theorem 3, it may, without loss of generality, be assumed that $a=-1$ and $b=1$. If $\phi(z)=(1+z)^{k-1}(1-z)^{l-1}$ and $P(z)=f'(z)/\phi(z)$, the latter being a polynomial of degree $\nu=n+1-k-l$, the requirements of the lemma just quoted will be satisfied and Theorem 3 will follow at once.

It will now be shown that *the envelope of the circles* (11) *when a and b vary independently in or on the unit circle is the circle of Theorem 2 with $p=k+l$.* It obviously suffices to find the envelope of the circles (11) when a and b vary on the unit circle. Every point of circle (11) may then have its coordinates written in the form

$$z = \frac{a+b}{2} + \theta\left(\frac{a-b}{2}\right)\cot\frac{\pi}{2(n+1-k-l)}$$

with $|\theta| \leqq 1$, and $|a| = |b|$. An angle ψ may be found so that either $a=be^{i\psi}$ or $b=ae^{i\psi}$ where $0 \leqq \psi \leqq \pi$. In either case

$$\begin{aligned} |z| &\leqq \cos\frac{\psi}{2} + \sin\frac{\psi}{2}\cot\frac{\pi}{2(n+1-k-l)} \\ &= \frac{\sin\left[\dfrac{\psi}{2} + \dfrac{\pi}{2(n+1-k-l)}\right]}{\sin\dfrac{\pi}{2(n+1-k-l)}} \leqq \csc\frac{\pi}{2(n+1-k-l)}. \end{aligned}$$

5. ***p*-valent polynomials.** An immediate corollary of Theorem 2 is the theorem:‡

* M. Fekete, Acta Litterarum ac Scientiarum, Szeged, vol. 1 (1923), pp. 98–100.

† M. Fekete, Mathematische Zeitschrift, vol. 22 (1925), p. 2, and Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 34 (1926), p. 211. See also M. Marden, Bulletin of the American Mathematical Society, vol. 38 (1932), p. 440; vol. 39 (1933), pp. 750–754.

‡ Alexander and Kakeya gave this theorem in the special case $p=1$. See the above references.

If the derivative of a polynomial $f(z)$ of degree $n \geqq 2$ has exactly $p-1$ zeros $(2 \leqq p < n)$ in the unit circle, then $f(z)$ has at most p zeros in or on the circle

$$|z| = \sin \frac{\pi}{2(n-p)} \cdot$$

For, if $f(z)$ had $p+1$ zeros in this circle, $f'(z)$ would have at least p zeros in or on the circle

$$|z| = \sin \frac{\pi}{2(n-p)} \csc \frac{\pi}{2(n-p)}$$

in contradiction to the hypothesis.

This corollary is essentially identical with the following theorem about p-valent polynomials:

THEOREM 4. *If the derivative of a polynomial $P(z)$ of degree n, $(n \geqq 2)$, has exactly $p-1$ zeros $(2 \leqq p < n)$ in or on the unit circle, then $P(z)$ is at most p-valent in or on the circle*

$$|z| = \sin \frac{\pi}{2(n-p)} \cdot$$

By a function's being p-valent in a given region R it is meant that the function takes on at least one value p times in R and no value more than p times in R. It suffices then merely to set $f(z) = P(z) - \gamma$, where γ is an arbitrary constant, in order to deduce Theorem 4 from the above corollary.

Finally, the same method of reasoning when used together with Theorem 2′ leads to the following more general conclusion giving a sufficient condition for a polynomial to be at most p-valent in a convex region K, not necessarily a circle.

THEOREM 4′. *Let K be a convex region and S the star-shaped region comprised of all points from which K subtends an angle of at least $\pi/(n-p)$ radians $(2 \leqq p < n)$. Then, if the derivative of any polynomial $P(z)$ of the nth degree has exactly $p-1$ zeros in S, the polynomial $P(z)$ is at most p-valent in K.*

UNIVERSITY OF WISCONSIN AT MILWAUKEE,
MILWAUKEE, WIS.