EXTREMAL POLYNOMIALS AND THE ZEROS OF THE DERIVATIVE OF A RATIONAL FUNCTION[1]

J. L. WALSH AND O. SHISHA

In the study of the zeros of extremal polynomials, Fekete and von Neumann indicated [1] that the derivative of an arbitrary polynomial $p(z) \equiv z^n + \cdots$ with simple zeros has extremal properties, and indeed is the polynomial $nz^{n-1} + \cdots$ which has the least suitably weighted Tchebycheff norm on the point set consisting of the zeros of $p(z)$. The object of the present note is to show that the zeros of the derivative of an arbitrary rational function are the zeros of a polynomial possessing analogous extremal properties.

THEOREM 1. *Let* $R(z) \equiv \prod_1^m (z-\alpha_j)^{m_j} \cdot \prod_1^n (z-\beta_j)^{-n_j}$ *be a rational function of* z *whose distinct finite zeros are the* α_j *of respective multiplicities* m_j *and whose distinct finite poles are the* β_j *of respective multiplicities* n_j. *Then the finite zeros of the derivative* $R'(z)$ *distinct from the* α_j *are the zeros of the extremal polynomial* $P(z) \equiv Nz^{m+n-1} + \cdots$, $N = \sum m_j - \sum n_j$, *required to assume the prescribed values* $P(\alpha_j) = m_j\omega'(\alpha_j)$ *in the points* α_j *and having minimum weighted Tchebycheff norm* $\max |P(\beta_j)/n_j\omega'(\beta_j)|$ *on the set* B: $\{\beta_1, \beta_2, \cdots, \beta_n\}$, *where* $\omega(z) \equiv \prod_1^m (z-\alpha_j) \cdot \prod_1^n (z-\beta_j)$.

The remark of Fekete and von Neumann was later extended [2] to include the case that $p(z)$ has multiple zeros, and we consider in Theorem 1 and below the analogue of the extension. In the proof of Theorem 1 we omit the case $m=0$ and the case $n=0$, essentially included in [1] and [2]; indeed an obvious modification of the present proof also includes these cases.

We identify the logarithmic derivative of $R(z)$:

$$\frac{R'(z)}{R(z)} \equiv \sum_1^m \frac{m_j}{z-\alpha_j} - \sum_1^n \frac{n_j}{z-\beta_j}, \tag{1}$$

after multiplication by $\omega(z)$, with the Lagrange interpolation formula for $P(z)$ and the point set consisting of the zeros of $\omega(z)$:

$$P(z) \equiv \omega(z) \sum_1^m \frac{P(\alpha_j)}{\omega'(\alpha_j)(z-\alpha_j)} + \omega(z) \sum_1^n \frac{P(\beta_j)}{\omega'(\beta_j)(z-\beta_j)}, \tag{2}$$

Presented to the Society, March 4, 1963; received by the editors May 1, 1963.

[1] Research supported (in part) by the Air Force Office of Scientific Research.

where $P(z)$ is the unique polynomial $Nz^{m+n-1}+\cdots$ which takes on the values $P(\alpha_j)$ in the points α_j and the values $P(\beta_j)$ in the points β_j,

$$P(\alpha_j) = m_j\omega'(\alpha_j), \qquad P(\beta_j) = -n_j\omega'(\beta_j). \tag{3}$$

It may be noted that any polynomial $Q(z)$ of the form $Nz^{m+n-1}+\cdots$, which takes the prescribed values $m_j\omega'(\alpha_j)$ in the points α_j, can be written

$$Q(z) \equiv \omega(z)\sum_1^m \frac{m_j}{z-\alpha_j} + \omega(z)\sum_1^n \frac{Q(\beta_j)}{\omega'(\beta_j)(z-\beta_j)}, \tag{4}$$

subject to the condition

$$N - \sum_1^m m_j = -\sum_1^n n_j = \sum_1^n \frac{Q(\beta_j)}{\omega'(\beta_j)} \qquad (<0). \tag{5}$$

For the Tchebycheff norm we introduce positive weights on B:

$$\mu_j = \mu(\beta_j) = 1/n_j\,|\,\omega'(\beta_j)\,|\,, \tag{6}$$

and in Theorem 1 the $Q(\beta_j)$ are to be chosen so that $(j=1, 2, \cdots, n)$

$$\max \mu_j\,|\,Q(\beta_j)\,| = \max \frac{|\,Q(\beta_j)\,|}{n_j\,|\,\omega'(\beta_j)\,|} \tag{7}$$

is as small as possible, subject to (5).

It will be convenient to have

LEMMA 1. *Given the weights* ν_j (>0); *the minimum of* $\max \nu_j|\lambda_j|$, $j=1, 2, \cdots, n$, *subject to the condition* $\sum\lambda_j=h$, *where* h (>0) *is preassigned and the (variable)* λ_j *are to be determined, is* $M_0=h/\sum(1/\nu_j)$, *given by* $\nu_j\lambda_j=M_0$ *for all* j.

We must have $\lambda_j\geqq 0$, for otherwise we may set $\lambda_j'=h|\lambda_j|/\sum|\lambda_k|$, whence $\sum\lambda_j'=h$, $|\lambda_j'|<|\lambda_j|$ for every j unless $\lambda_j=0$. With $\lambda_j\geqq 0$ and $M_0=\max \nu_j\lambda_j$, we have $\lambda_j\leqq M_0\sum(1/\nu_j)$, so either we have $\nu_j\lambda_j=M_0$ for every j, $M_0=h/\sum(1/\nu_j)$, or for some j we have $\nu_j\lambda_j<M_0$, $M_0>h/\sum(1/\nu_j)$. Minimum M_0 occurs for $\nu_j\lambda_j=h/\sum(1/\nu_k)$ for every j. Of course multiplication of all ν_j by a positive constant does not alter the extremal λ_j.

In the situation of Theorem 1 we set $\lambda_j=-P(\beta_j)/\omega'(\beta_j)$, $\nu_j=1/n_j$, $\mu_j=1/n_j|\omega'(\beta_j)|$, whence $\mu_j|P(\beta_j)|=\nu_j|\lambda_j|$. When the members of (7) are a minimum, each of the n elements $\nu_j\lambda_j$ should by Lemma 1 be unity; this condition is both necessary and sufficient that (7) be a minimum, so the unique extremal polynomial is $P(z)$ as defined by (2) and (3); Theorem 1 is established. The prescribed values of the

Tchebycheff polynomial in the α_j are $P(\alpha_j)=m_j\omega'(\alpha_j)$, and the prescribed weights in the points β_j are given by (6). The extremal values in the β_j are $P(\beta_j)=-n_j\omega'(\beta_j)$.

Naturally the roles of the α_j and the β_j may be reversed in this theorem; we may prescribe values of the polynomial $P(z)$ in the β_j and determine the least norm on $\{\alpha_j\}$.

The proof of Theorem 1 is naturally related to the proofs given in [**1**] and [**2**], and is also related to the study of restricted infrapolynomials in [**5**]. The present emphasis is on restricted extremal polynomials, and the direct use of Lagrange's formula as in (2) seems here more favorable than the methods of [**5**].

If it is desired to describe the extremal properties not of the polynomial $P(z)$ defined in Theorem 1, but of a polynomial whose zeros are *all* the finite zeros of $R'(z)$ and with the same multiplicities, we can no longer use the simple Lagrange formula (2); however, it is sufficient to set

$$P_1(z) \equiv u_1(z)\cdot P(z), \qquad u_1(z) \equiv \prod_1^m (z-\alpha_j)^{m_j-1}. \tag{8}$$

In the neighborhood of α_j we have the prescription

$$P_1(z) \equiv m_j(z-\alpha_j)^{m_j-1} + E_j(z), \tag{9}$$

where E_j is a polynomial with the factor $(z-\alpha_j)^{m_j}$. Moreover for $P_1(z)$ we prescribe in the points β_j the new weights

$$\mu_j' = \mu_j / \left| u_1(\beta_j) \right| . \tag{10}$$

Then $P_1(z)$ is a polynomial $Nz^M+\cdots$, $M=\sum m_j+n-1$, whose prescribed values in the simple zeros α_j of $R(z)$ are given by $P_1(\alpha_j)=u_1(\alpha_j)\cdot P(\alpha_j)$, and whose behavior in the multiple zeros of $R(z)$ is indicated by (9); consequently $P_1(z)$ is the unique thus restricted polynomial of least Tchebycheff norm on B with weights μ_j' as in (10).

Theorem 1 expresses the extremal polynomial $P(z)\equiv\omega(z)R'(z)/R(z)$ as a restricted (i.e., $P(z)\equiv Nz^{m+n-1}+\cdots$ with $P(\alpha_j)=m_j\omega'(\alpha_j)$) polynomial of least Tchebycheff norm on B. For every p ($\geqq 1$) and with suitably chosen weights in the β_j, *this extremal Tchebycheff polynomial is also a similarly restricted polynomial of least pth power norm on B*, namely minimizing the norm

$$\sum_{j=1}^n \frac{\left| P(\beta_j) \right|^p}{(n_j \left| \omega'(\beta_j) \right|)^{p-1}} ;$$

the extremal polynomial is unique for $p>1$. The proof of this statement is closely analogous to the proof in [3, p. 373], and details are left to the reader.

The remark of Fekete and von Neumann already mentioned is especially appropriate in the study of the geometry of zeros, for (as they indicate) the classical theorem of Lucas asserting that the zeros of $p'(z)$ lie in the convex hull of the zeros of $p(z)$ follows from the theorem of Fejér that the zeros of a polynomial of minimum norm (indeed, of any infrapolynomial) on a point set lie in the convex hull of that set—assuming the set to be sufficiently numerous. Theorem 1 enables one similarly to prove Bôcher's theorem (if two disjoint circular regions contain respectively the zeros and poles of a rational function they contain all its finite critical points) from results on restricted extremal polynomials, and similarly to prove various other propositions concerning zeros. Compare Theorem 3 of [5].

Here several related remarks are appropriate.

1°. If a rational function can be written in the form

$$\sum_1^m \frac{a_j}{z-\alpha_j} - \sum_1^n \frac{b_j}{z-\beta_j}, \qquad a_j>0,\ b_j>0,\ \sum a_j = \sum b_j, \tag{11}$$

and if two disjoint circular regions contain respectively the α_j and β_j, then these regions contain all finite zeros of (11)—the proof is essentially that of Bôcher's theorem [4, §4.2]. If we have (11) except that now $\sum a_j=\mu\neq\sum b_j=\nu$, and if the α_j lie in the disk $|z-\alpha_0|\leqq r_1$ and the β_j in the disk $|z-\beta_0|\leqq r_2$, then [4, §4.2.4] all finite zeros of (11) exterior to those disks lie in the disk

$$\left| z - \frac{\mu\beta_0 - \nu\alpha_0}{\mu-\nu} \right| \leqq \frac{\mu r_2 + \nu r_1}{|\mu-\nu|}.$$

Numerous further results follow at once as in [4, Chapters IV, V], for instance if the α_j and β_j are real, or if (11) admits other symmetries.

2°. If we replace (11) by

$$\sum_1^m \frac{a_j}{z-\alpha_j}, \qquad a_j>0, \tag{12}$$

various results can be established as in [4, Chapters I–III], including for instance the analogue of Lucas's theorem, that all zeros of (12) lie in the convex hull of the α_j.

3°. If we consider (generalization of (11))

$$(13)\qquad \epsilon\sum_1^m \frac{a_j}{z-\alpha_j}+\sum_1^n \frac{b_j}{z-\beta_j}, \qquad a_j>0,\ b_j>0,\ |\epsilon|=1,$$

and if the α_j and β_j lie in the respective disks $|z-\alpha_0|\leqq r_1$, $|z-\beta_0|\leqq r_2$, then all zeros of (13) exterior to those disks lie in the disk (provided $\epsilon\mu+\nu\neq 0$)

$$\left| z-\frac{\epsilon\mu\beta_0+\nu\alpha_0}{\epsilon\mu+\nu}\right| \leqq \frac{\mu r_2+\nu r_1}{|\epsilon\mu+\nu|}, \qquad \mu=\sum a_j,\ \nu=\sum b_j.$$

The significance of (11), (12), and (13), when their product with $\omega(z)$ is identified with the general Lagrange interpolation formula as in the second member of (2), is that in (11) and (12) we have $\arg P(\alpha_j)=\arg \omega'(\alpha_j)$; in (11) we have $\arg P(\beta_j)=\arg[-\omega'(\beta_j)]$; in (13) we have $\arg P(\alpha_j)=\arg[\epsilon\omega'(\alpha_j)]$ and $\arg P(\beta_j)=\arg\omega'(\beta_j)$.

4°. Certain extremal properties lead directly to functions of form (11), (12), or (13); *and conversely*, each function of such a form (after multiplication by $\omega(z)$) can be interpreted as a polynomial characterized by suitable extremal properties, as in Theorem 1. If a polynomial $P(z)$ has the form (2) where $\alpha_j\neq\beta_k$, if the numbers $P(\alpha_j)/\omega'(\alpha_j)$ are all prescribed and positive, and if the numbers $P(\beta_k)$ are to be determined so that $P(z)\equiv Az^{m+n-1}+\cdots$ where A is fixed and $A<\sum P(\alpha_j)/\omega'(\alpha_j)$, and so that $P(z)$ is a restricted infrapolynomial (or more particularly a polynomial of least norm) on the set B, then we must have $P(\beta_j)/\omega'(\beta_j)<0$ for every j, $P(z)/\omega(z)$ is of form (11), and 1° applies. Compare here [5], where such infrapolynomials are considered in some detail; compare also [6]. However, if the $P(\alpha_j)/\omega'(\alpha_j)$ are all prescribed and positive, and if A is fixed with $A>\sum P(\alpha_j)/\omega'(\alpha_j)$, then for the infrapolynomial $P(z)$ of (2) we must have $P(\beta_j)/\omega'(\beta_j)>0$ for every j, and $P(z)/\omega(z)$ is of form (12); 2° applies. If we have $A=\sum P(\alpha_j)/\omega'(\alpha_j)$, then for the infrapolynomial we must have $P(\beta_j)=0$ for every j; $P(z)/\omega(z)$ is of form (12) and 2° applies. If $P(z)$ has the form (2) where $\alpha_j\neq\beta_k$, if the numbers $P(\alpha_j)=\epsilon a_j\omega'(\alpha_j)$ are prescribed and the $P(\alpha_j)/\omega'(\alpha_j)$ all have the same argument, and if the $P(\beta_j)$ are to be determined so that $P(z)\equiv(\epsilon\mu+\nu)z^{m+n-1}+\cdots$, $\epsilon\mu+\nu\neq 0$, $\nu>0$, is a restricted infrapolynomial (or a polynomial of least norm) on B; then we must have $P(\beta_j)/\omega'(\beta_j)>0$ for every j, and $P(z)/\omega(z)$ is of form (13). Remark 3° applies.

Beyond propositions 1°, 2°, and 3°, and identification of the polynomials $P(z)$ with restricted infrapolynomials, and conversely, there can be proved results of a different nature (compare [5] and [6]) especially concerning point sets and polynomials possessing various symmetries. Moreover the original functions need not be rational; for instance in Theorem 1 the m_j (>0) and n_j (>0) need not be rational.

References

1. M. Fekete and J. von Neumann, *Über die Lage der Nullstellen gewisser Minimumpolynome*, Jber. Deutsch. Math.-Verein. **31** (1922), 125–138.

2. T. S. Motzkin and J. L. Walsh, *On the derivative of a polynomial and Chebyshev approximation*, Proc. Amer. Math. Soc. **4** (1953), 76–87.

3. ———, *The least pth power polynomials on a finite point set*, Trans. Amer. Math. Soc. **83** (1956), 371–396.

4. J. L. Walsh, *The location of critical points*, Amer. Math. Soc. Colloq. Publ. Vol. 34, Amer. Math. Soc., Providence, R. I., 1950.

5. J. L. Walsh and O. Shisha, *The zeros of infrapolynomials with prescribed values at given points*, Proc. Amer. Math. Soc. **14** (1963), 839–844.

6. J. L. Walsh, *Geometry of the zeros of the sums of linear fractions*, Trans. Amer. Math. Soc. (to appear).

Harvard University and
Aerospace Research Laboratories,
Wright-Patterson Air Force Base

NONLINEAR DIFFERENTIAL EQUATIONS WITH FORCING TERMS[1]

FRED BRAUER

1. In this paper, we shall study the solutions of a differential equation containing a linear term with constant coefficients, a nonlinear term, and a forcing term depending only on the independent variable. We shall attempt to compare these with solutions of the equation obtained by neglecting the nonlinear term. This is a problem which frequently arises in physical examples, where the linear equation is solved and its solution is used to describe approximately the motion governed by the nonlinear equation. We shall see that the solutions of the two equations do behave similarly if the nonlinear term is small enough. This does not settle the question by any means, as in practice the nonlinear term is frequently not small enough for our result to be applicable.

Received by the editors June 14, 1963.

[1] This research was supported in part by the National Science Foundation (G-24355).