EXTREMAL POLYNOMIALS AND THE ZEROS OF THE DERIVATIVE OF A RATIONAL FUNCTION

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 \( p(z) = z^n + \cdots \) 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 \( R(z) = \prod_{1}^{m} (z - \alpha_j)^{m_j} \cdot \prod_{1}^{n} (z - \beta_j)^{-n_j} \) be a rational function of \( z \) whose distinct finite zeros are the \( \alpha_j \) of respective multiplicities \( m_j \) and whose distinct finite poles are the \( \beta_j \) of respective multiplicities \( n_j \). Then the finite zeros of the derivative \( R'(z) \) distinct from the \( \alpha_j \) are the zeros of the extremal polynomial \( P(z) = N z^{m+n} + \cdots \), \( N = \sum m_j - \sum n_j \), required to assume the prescribed values \( P(\alpha_j) = m \omega(\alpha_j) \) in the points \( \alpha_j \) and having minimum weighted Tchebycheff norm

\[
\max \left\{ \frac{P(\beta_j)}{n \omega'(\beta_j)} \right\} \text{ on the set } B: \{ \beta_1, \beta_2, \cdots, \beta_n \}, \text{ where } \omega(z) = \prod_{1}^{m} (z - \alpha_j) \cdot \prod_{1}^{n} (z - \beta_j).
\]

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 \( R(z) \):

\[
\frac{R'(z)}{R(z)} = \sum_{1}^{m} \frac{m_j}{z - \alpha_j} - \sum_{1}^{n} \frac{n_j}{z - \beta_j},
\]

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

\[
P(z) = \omega(z) \sum_{1}^{m} \frac{P(\alpha_j)}{\omega'(\alpha_j)(z - \alpha_j)} + \omega(z) \sum_{1}^{n} \frac{P(\beta_j)}{\omega'(\beta_j)(z - \beta_j)},
\]

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 $\alpha_j$ and the values $P(\beta_j)$ in the points $\beta_j$,

$$P(\alpha_j) = m_j \omega'(\alpha_j), \quad P(\beta_j) = -n_j \omega'(\beta_j).$$

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 $\alpha_j$, can be written

$$Q(z) = \omega(z) \sum_{j=1}^{m} \frac{m_j}{z - \alpha_j} + \omega(z) \sum_{j=1}^{n} \frac{Q(\beta_j)}{\omega'(\beta_j)(z - \beta_j)},$$

subject to the condition

$$N - \sum_{j=1}^{m} m_j = -\sum_{j=1}^{n} n_j = -\sum_{j=1}^{n} \frac{Q(\beta_j)}{\omega'(\beta_j)} \quad (<0).$$

For the Tchebycheff norm we introduce positive weights on $B$:

$$\mu_j = \mu(\beta_j) = 1/n_j \left| \omega'(\beta_j) \right|,$$

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

$$\max_{j=1, 2, \cdots, n} \left| Q(\beta_j) \right| = \max_{j=1, 2, \cdots, n} \frac{\left| Q(\beta_j) \right|}{n_j \left| \omega'(\beta_j) \right|}$$

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

It will be convenient to have

**Lemma 1.** Given the weights $\nu_j$ ($>0$); the minimum of $\max_{j=1, 2, \cdots, n} \nu_j |\lambda_j|$, $j=1, 2, \cdots, n$, subject to the condition $\sum \lambda_j = h$, where $h$ ($>0$) is pre-assigned and the (variable) $\lambda_j$ 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 $\lambda_j \geq 0$, 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 $\lambda_j \geq 0$ and $M_0 = \max \nu_j \lambda_j$, we have $\lambda_j \leq M_0 \sum (1/\nu_j)$, 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 $\nu_j$ by a positive constant does not alter the extremal $\lambda_j$.

In the situation of Theorem 1 we set $\lambda_j = -P(\beta_j)/\omega'(\beta_j)$, $\nu_j = 1/n_j$, $\mu_j = 1/n_j \left| \omega'(\beta_j) \right|$, 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 \( \alpha_j \) are \( P(\alpha_j) = m_j \omega'(\alpha_j) \), and the prescribed weights in the points \( \beta_j \) are given by (6). The extremal values in the \( \beta_j \) are \( P(\beta_j) = -n_j \omega'(\beta_j) \).

Naturally the roles of the \( \alpha_j \) and the \( \beta_j \) may be reversed in this theorem; we may prescribe values of the polynomial \( P(z) \) in the \( \beta_j \) 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 infra-polynomials 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

\[
P_i(z) = u_1(z) - P(z), \quad u_1(z) = \prod_{j=1}^n (z - \alpha_j)^{m_j-1}.
\]

In the neighborhood of \( \alpha_j \) we have the prescription

\[
P_1(z) = m_j (z - \alpha_j)^{m_j-1} + E_j(z),
\]

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

\[
\mu_j' = \frac{\mu_j}{u_1(\beta_j)}.
\]

Then \( P_1(z) \) is a polynomial \( Nz^m + \cdots, M = \sum m_j + n - 1 \), whose prescribed values in the simple zeros \( \alpha_j \) 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 \( \mu_j' \) as in (10).

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

\[
\sum_{j=1}^n \left( \frac{|P(\beta_j)|^p}{n_j |\omega'(\beta_j)|^{p-1}} \right)^{1/p}.
\]
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

\[
\sum_{1}^{m} \frac{a_{j}}{z - \alpha_{j}} - \sum_{1}^{n} \frac{b_{j}}{z - \beta_{j}}, \quad a_{j} > 0, b_{j} > 0, \sum a_{j} = \sum b_{j},
\]

and if two disjoint circular regions contain respectively the \( \alpha_{j} \) and \( \beta_{j} \), 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 \( \alpha_{j} \) lie in the disk \( |z - \alpha_{0}| \leq r_{1} \) and the \( \beta_{j} \) in the disk \( |z - \beta_{0}| \leq r_{2} \), then [4, §4.2.4] all finite zeros of (11) exterior to those disks lie in the disk

\[
\left| z - \frac{\mu \beta_{0} - \nu \alpha_{0}}{\mu - \nu} \right| \leq \frac{\mu r_{2} + \nu r_{1}}{|\mu - \nu|}.
\]

Numerous further results follow at once as in [4, Chapters IV, V], for instance if the \( \alpha_{j} \) and \( \beta_{j} \) are real, or if (11) admits other symmetries.

2°. If we replace (11) by

\[
\sum_{1}^{m} \frac{a_{j}}{z - \alpha_{j}}, \quad a_{j} > 0,
\]

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 \( \alpha_{j} \).

3°. If we consider (generalization of (11))
and if the \( \alpha_j \) and \( \beta_j \) lie in the respective disks \( |z - \alpha_0| \leq r_1, |z - \beta_0| \leq r_2 \), then all zeros of (13) exterior to those disks lie in the disk (provided \( \epsilon \mu + \nu \neq 0 \))

\[
\left| z - \frac{\epsilon \mu \beta_0 + \nu \alpha_0}{\epsilon \mu + \nu} \right| \leq \frac{\mu r_2 + \nu r_1}{\epsilon \mu + \nu}, \quad \mu = \sum a_j, \nu = \sum b_j.
\]

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 \( \arg P(\alpha_j) = \arg [\omega'(\alpha_j)] \) 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 \( P(z) = Az^{m+n-1} + \cdots \) 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 \( P(\alpha_j) = \epsilon \omega'(\alpha_j) \) 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 \( P(z) = (\epsilon \mu + \nu)z^{m+n-1} + \cdots, \epsilon \mu + \nu \neq 0, \nu > 0 \), 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 $m_i (>0)$ and $n_i (>0)$ need not be rational.

References


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

NONLINEAR DIFFERENTIAL EQUATIONS
WITH FORCING TERMS

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).