A REFINEMENT OF PELLET'S THEOREM

MORRIS MARDEN

1. Introduction. S. Lipka¹ has recently announced a refinement of the classic theorem of Cauchy that all the zeros of the polynomial

(1.1) \( f(z) = a_0 + a_1z + \cdots + a_nz^n, \quad a_n \neq 0, \)

lie in the circle \( |z| \leq r \), where \( r \) is the positive root of the real equation

(1.2) \( F_n(z) = |a_0| + |a_1| z + \cdots + |a_{n-1}| z^{n-1} - |a_n| z^n = 0. \)

Lipka's refinement consists in replacing the circle \( |z| = r \) by a curve \( G(r_0, r_n; n, \alpha_0) \) which bounds a gear-wheel region. This region is formed by deleting from the circle \( |z| \leq r \) the points common to the annular ring \( 0 < r_0 < |z| \leq r \) and to the \( n \) sectors

(1.3) \( \frac{\alpha_0}{n} - \frac{\pi}{2n} + \frac{2\pi k}{n} \leq \arg z \leq \frac{\alpha_0}{n} + \frac{\pi}{2n} + \frac{2\pi k}{n}, \quad k = 0, 1, \cdots, n-1. \)

In these formulas \( r_0 \) is the positive root of the equation

(1.4) \( \Phi_n(z) = |a_1| + |a_2| z + \cdots + |a_{n-1}| z^{n-2} - |a_n| z^{n-1} = 0 \)

and \( \alpha_0 = \arg a_0/a_n. \)

Now, the Cauchy theorem is but a special case of the following theorem due to Pellet.²

PELLET'S THEOREM. If the polynomial

(1.5) \( f(z) = a_0 + a_1z + \cdots + a_pz^p + \cdots + a_nz^n, \quad a_p \neq 0, \)

is such that the real polynomial

(1.6) \( F_p(z) = |a_0| + |a_1| z + \cdots + |a_{p-1}| z^{p-1} - |a_p| z^p + |a_{p+1}| z^{p+1} + \cdots + |a_n| z^n \)

Presented to the Society, September 3, 1947; received by the editors August 22, 1947.

¹ S. Lipka, Monatsh. für Mathematik und Physik vol. 50 (1944) pp. 209–221.
has two positive zeros $r$ and $R$ with $r < R$, then $f(z)$ has exactly $p$ zeros in or on the circle $|z| \leq r$ and no zeros in the annular ring $r < |z| < R$.

It is Pellet's theorem which we propose to refine as indicated in the following theorem.

**Theorem 1.1.** Under the hypotheses of Pellet's Theorem the polynomial

\[
\Phi_p(z) = |a_1| + |a_2| z + \cdots + |a_{p-1}| z^{p-2} - |a_p| z^{p-1} \\
+ |a_{p+1}| z^p + \cdots + |a_n| z^{n-1}
\]

has also two positive zeros $r_0$ and $R_0$ with

\[
r_0 < r < R < R_0.
\]

Furthermore, $f(z)$ has exactly $p$ zeros in or on the curve $G(r_0, r; p, \alpha_0)$ where $\alpha_0 = \arg a_0/a_p$ and no zeros between the curves $G(r_0, r; p, \alpha_0)$ and $G(R, R_0; p, \alpha_0 + \pi)$.

Theorem 1.1 will be proved in §2 and applied in §3 to the refinement of various known bounds on the zeros of a polynomial. Finally, the theorem will be generalized in §4, first by replacing the polynomial $\Phi_p(z)$ by the polynomial $\Phi_kp(z) = F_p(z) - |a_k| z^k$ and secondly by replacing the polynomial $f(z)$ by a power series.

2. Proof of Theorem 1.1. Let us first prove the existence of the roots $r_0$ and $R_0$ of equation $\Phi_p(z) = 0$ and the validity of inequality (1.8). Since $r$ and $R$ are the positive zeros of $F_p(z)$, it follows from (1.6) that, for any sufficiently small positive number $\epsilon$,

\[
F_p(\rho) < 0 \quad \text{if } r + \epsilon \leq \rho \leq R - \epsilon.
\]

In view of the equation

\[
F_p(z) = |a_0| + z \Phi_p(z),
\]

the zeros $r$ and $R$ of $F_p(z)$ satisfy the relations

\[
\Phi_p(r) = -|a_0| /r < 0, \quad \Phi_p(R) = -|a_0| /R < 0.
\]

When taken together with the facts that

\[
\Phi_p(0) > 0, \quad \Phi_p(\infty) > 0,
\]

the relations (2.3) imply the existence of two positive zeros $r_0$ and $R_0$ of $\Phi_p(z)$ and the validity of inequality (1.8), as well as the inequality

\[
\Phi_p(\rho) < 0
\]

for $r_0 + \epsilon \leq \rho \leq R_0 - \epsilon$. 
Let us now set \( z = \rho e^{i\theta} \) and
\[
(2.6) \quad a_k/a_p = A_k e^{a k_i}, \quad k = 0, 1, \ldots, n.
\]
In this notation, the real part of \( \rho f(z)/a_p z^p \) is
\[
\text{Re} \left[ \frac{\rho f(z)}{a_p z^p} \right] = \sum_{i=0}^{p-1} A_i \rho^i \cos [(p - j)\theta - \alpha_i] + \rho^p
\]
\[
+ \sum_{j=p+1}^{n} A_j \rho^i \cos [(j - p)\theta + \alpha_i]
\]
and the inequalities (2.1) and (2.5) become
\[
\text{(2.8)} \quad \rho^p > A_0 + A_1 \rho + \cdots + A_{p-1} \rho^{p-1} + A_{p+1} \rho^{p+1} + \cdots + A_n \rho^n
\]
for \( r + \epsilon \leq \rho \leq R - \epsilon \), and
\[
\text{(2.9)} \quad \rho^p > A_0 \rho + A_2 \rho^2 + \cdots + A_{p-1} \rho^{p-1} + A_{p+1} \rho^{p+1} + \cdots + A_n \rho^n
\]
for \( r_0 + \epsilon \leq \rho \leq R_0 - \epsilon \).

On substituting from inequality (2.8) into (2.7), we find
\[
\text{Re} \left[ \frac{\rho f(z)}{a_p z^p} \right] > A_0 + \sum_{i=1}^{p-1} A_i \rho^i \cos [(p - j)\theta - \alpha_i] + 1
\]
\[
+ \sum_{j=p+1}^{n} A_j \rho^i \cos [(j - p)\theta + \alpha_i] + 1 \geq 0
\]
for \( r + \epsilon \leq \rho \leq R - \epsilon \). On substituting from inequality (2.9) into (2.7), we find
\[
\text{Re} \left[ \frac{\rho f(z)}{a_p z^p} \right] > A_0 \cos (p\theta - \alpha_0)
\]
\[
+ \sum_{i=1}^{p-1} A_i \rho^i \cos [(p - j)\theta - \alpha_i] + 1
\]
\[
+ \sum_{j=p+1}^{n} A_j \rho^i \cos [(j - p)\theta + \alpha_i] + 1
\]
for \( r_0 + \epsilon \leq \rho \leq R_0 - \epsilon \). The right side of (2.11) is surely non-negative if \( \theta \) is such that \( \cos (p\theta - \alpha_0) \geq 0 \), that is, such that
\[
-\frac{\pi}{2} + 2\pi k \leq p\theta - \alpha_0 \leq \frac{\pi}{2} + 2\pi k,
\]
where \( k \) is an integer; that is, if
\[
\frac{\alpha_0}{p} - \frac{\pi}{2p} \leq \frac{2\pi k}{p} \leq \frac{\alpha_0}{p} + \frac{\pi}{2p} + \frac{2\pi k}{p}, \quad k = 0, 1, \ldots, p.
\]
In other words,

\[(2.12) \quad \text{Re} \left( \frac{p^2 f(z)}{a_z z^p} \right) > 0 \]

and hence \( f(z) \neq 0 \) at all points \( z \) between the curves \( G(r_0, r; \rho, \alpha_0) \) and \( G(R, R_0; \rho, \alpha_0 + \pi) \).

Inequality (2.12) also may be used to show that in or on the curve \( G(r_0, r; \rho, \alpha_0) \), there are exactly \( p \) zeros of \( f(z) \). For, let us consider the net change \( \Delta g \arg w \) in the argument of the point \( w = \left[ \frac{p^2 f(z)}{a_z z^p} \right] \) as \( z \) describes counterclockwise the curve \( G_* = G(r_0 + \epsilon, r + \epsilon; \rho, \alpha_0) \) where \( \epsilon \) is a small positive number. Since \( \text{Re} \ (w) > 0 \), \( w \) describes a closed curve entirely in the right-half \( w \)-plane. That is, \( \Delta g \arg w = 0 \) on this curve. This means that the function \( w \) has as many zeros as poles in the curve \( G_* \) and this, in turn, means that \( f(z) \) has precisely \( p \) zeros in \( G_* \) for every sufficiently small positive \( \epsilon \).

3. Applications. Let us first apply Theorem 1.1 to the class of polynomials

\[(3.1) \quad f(z) = b_0 e^{i\theta_0} + (b_1 - b_0) e^{i\theta_1} z + \cdots + (b_m - b_{m-1}) e^{i\theta_m} z^m - b_m e^{i\theta_{m+1}} z^{m+1} \]

where the \( b_j \) are real numbers such that

\[(3.2) \quad b_{p-1} < b_{p-2} < \cdots < b_0 < 0 < b_m < b_{m-1} < \cdots < b_p. \]

The corresponding polynomials \( F_p(z) \) and \( \Phi_p(z) \) are

\[(3.3) \quad F_p(z) = -b_0 + (b_0 - b_1) z + \cdots + (b_{p-2} - b_{p-1}) z^{p-1} + (b_p - b_{p-1}) z^p + \cdots + (b_{m-1} - b_m) z^m + b_m z^{m+1}, \]

\[(3.4) \quad \Phi_p(z) = (b_0 - b_1) + \cdots + (b_{p-2} - b_{p-1}) z^{p-2} + (b_p - b_{p-1}) z^{p-1} + \cdots + (b_{m-1} - b_m) z^{m-1} + b_m z^m. \]

On defining

\[(3.5) \quad g(z) = b_0 + b_1 z + \cdots + b_m z^m, \]

we may write

\[ F_p(z) = (z - 1) g(z), \quad z \Phi_p(z) = b_0 + g(z)(z - 1). \]

Clearly \( F_p(1) = 0 \). Since \( F_p(1 + \delta) = \delta g(1 + \delta) \), then for \( \delta \) sufficiently small \( g(1) > 0 \) implies that \( F_p(1 + \delta) > 0 \) or \( < 0 \) according as \( \delta > 0 \) or \( < 0 \) and \( g(1) < 0 \) implies that \( F_p(1 + \delta) < 0 \) or \( > 0 \) according as \( \delta > 0 \) or
<0. That is, using the notation of Theorem 1.1, we see that

\[ r_0 < r < 1 = R < R_0 \quad \text{if } g(1) > 0, \]
\[ r_0 < r = 1 < R < R_0 \quad \text{if } g(1) < 0, \]
\[ \alpha_0 = \beta_0 - \beta_p + \pi. \]

We thereby conclude that the following is true.

**Theorem 3.1.** Let \( f(z), \Phi_p(z) \) and \( g(z) \) denote the polynomials (3.1), (3.4) and (3.5) respectively. Then, if \( g(1) > 0 \), \( f(z) \) has exactly \( p \) zeros in the curve \( G(r_0, 1; p, \beta_0 - \beta_p + \pi) \) and \( g(z) \) has \( p \) zeros in the curve \( G(r_0, 1; p, \pi) \). If \( g(1) < 0 \), \( f(z) \) has \( p \) zeros in or on the curve \( G(r_0, 1; p, \beta_0 - \beta_p + \pi) \) and \( g(z) \) has \( p - 1 \) zeros in or on the curve \( G(r_0, 1; p, \pi) \).

An analogous result for \( g(z) \) with, however, curve \( G(r_0, 1; p, \pi) \) replaced by the circle \( |z| = 1 \) was first stated by Berwald.\(^3\) His result was a generalization of the Kakeya-Eneström\(^4\) theorem that all the zeros of the real polynomial (3.5) with \( 0 < b_0 < b_1 < \cdots < b_n \) lie in or on the unit circle \( |z| = 1 \). Our analogy to the Kakeya-Eneström theorem will be included in the following theorem.

**Theorem 3.2.** Every polynomial of the form

\[ f(z) = \sum_{j=0}^{n} (b_j - b_{j-1})e^{\theta_j}z^j, \quad b_{-1} = b_n = 0 < b_0 < b_1 < \cdots < b_{n-1}, \]

has all of its zeros in or on the curve \( G(r_0, 1; n, \beta_0 - \beta_n + \pi) \) where \( r_0 \) is the positive root of the equation

\[ \Phi_n = (b_1 - b_0) + (b_2 - b_1)z + \cdots + (b_{n-1} - b_{n-2})z^{n-2} - b_{n-1}z^{n-1} = 0. \]

Furthermore, every polynomial of the form

\[ g(z) = b_0 + b_1z + \cdots + b_{n-1}z^{n-1}, \quad 0 < b_0 < b_1 < \cdots < b_{n-1}, \]

has all of its zeros in or on the curve \( G(r_0, 1; n, \pi) \).

This theorem may be derived from Theorem 3.1 indirectly by a limiting process or directly by the same methods as used for Theorem 3.1.

In our next application, we shall use Theorem 1.1 just in the case \( p = n \). This restriction is made only to simplify the statement of results, since a similar application may be made when \( p \) is an arbitrary integer, \( 0 < p \leq n \). The result to be proved is the following.


Theorem 3.3. Let \( \lambda_1, \lambda_2, \cdots, \lambda_n \) and \( \mu_1, \mu_2, \cdots, \mu_{n-1} \) be any two sets of numbers such that

\[
\sum_{i=1}^{n} \frac{1}{\lambda_i} = 1, \quad \sum_{j=1}^{n-1} \frac{1}{\mu_j} = 1; \quad 0 < \mu_j \leq \lambda_i, \ j = 1, 2, \cdots, n - 1.
\]

For the polynomial \( f(z) = a_0 + a_1 z + \cdots + a_n z^n \), let

\[
M = \max \left[ \lambda_k \left| a_{n-k} \right| / \left| a_n \right| \right]^{1/k}, \quad k = 1, 2, \cdots, n,
\]

(3.6) \[M_0 = \max \left[ \mu_k \left| a_{n-k} \right| / \left| a_n \right| \right]^{1/k}, \quad k = 1, 2, \cdots, n - 1.
\]

Then all the zeros of \( f(z) \) lie in or on the curve \( G(M_0, M; n, \alpha_0) \), where \( \alpha_0 = \arg (a_0/a_n) \).

From (3.6) and (3.7), obviously \( 0 < M_0 < M \). Also,

\[\lambda_k \left| a_{n-k} \right| \leq \left| a_n \right| M^k, \quad \mu_k \left| a_{n-k} \right| \leq \left| a_n \right| M_0^k\]

and thus

\[
\sum_{k=1}^{n} \left| a_{n-k} \right| M^{n-k} \leq \sum_{k=1}^{n} \left( \frac{1}{\lambda_k} \right) \left| a_n \right| M^n = \left| a_n \right| M^n,
\]

(3.8)

\[
\sum_{k=1}^{n-1} \left| a_{n-k} \right| M_0^{n-k} \leq \sum_{k=1}^{n-1} \left( \frac{1}{\mu_k} \right) \left| a_n \right| M_0^n = \left| a_n \right| M_0^n.
\]

(3.9)

An equality in (3.8) would imply that \( M \) is the positive root \( r \) of the equation (1.2) whereas an inequality in (3.8) would imply that \( M > r \). Likewise, an equality in (3.9) would imply that \( M_0 \) is the positive root \( r_0 \) of the equation (1.4) whereas an inequality in (3.9) would imply that \( M_0 > r_0 \). Since by Theorem 1.1 all the zeros of \( f(z) \) lie in or on the curve \( G(r_0, r; n, \alpha_0) \), they surely all lie in or on the curve \( G(M_0, M; n, \alpha_0) \).

Theorem 3.3 whose proof we have just completed is a refinement of the result due to Fujiwara\(^6\) that all the zeros of \( f(z) \) lie in or on the circle \( |z| \leq M \).

As a simple application of Theorem 3.3, let us take \( \lambda_j = n \) for \( j = 1, 2, \cdots, n \) and \( \mu_j = n-1 \) for \( j = 1, 2, \cdots, n-1 \). We obtain thereby the following corollary.

Corollary 3.3a. For the polynomial \( f(z) = a_0 + a_1 z + \cdots + a_n z^n \) let \( N = \max \left[ \left| a_{n-k} \right| / \left| a_n \right| \right]^{1/k}, \ k = 1, 2, \cdots, n \), and \( N_0 = \max \left[ (n-1)^k \left| a_{n-k} \right| / a_n \right]^{1/k}, k = 1, 2, \cdots, n-1 \). Then all the zeros of \( f(z) \) lie in or on the curve \( G(N_0, N; n, \alpha_0) \) where \( \alpha_0 = \arg (a_0/a_n) \).

As another simple application of Theorem 3.3, let us take
\[
\lambda_k = \sum_{j=0}^{n-1} \left| \frac{a_j}{a_n} \right|, \quad k = 0, 1, 2, 3, \ldots, n,
\]
\[
\mu_k = \sum_{j=1}^{n-1} \left| \frac{a_j}{a_n} \right|, \quad k = 0, 1, 2, \ldots, n - 1.
\]
Clearly,
\[
\sum_{j=1}^{n} 1/\lambda_k = 1, \quad \sum_{j=1}^{n-1} 1/\mu_j = 1.
\]
Here
\[
M = \max \left[ \left( \sum_{j=0}^{n-1} \left| \frac{a_j}{a_n} \right| \right)^{1/k} \right] = \lambda_0 \text{ or } \lambda_0^{1/n}
\]
according as \( \lambda_0 > 1 \) or \( < 1 \), and
\[
M_0 = \max \left[ \left( \sum_{j=1}^{n-1} \left| \frac{a_j}{a_n} \right| \right)^{1/k} \right] = \mu_0 \text{ or } \mu_0^{1/n}
\]
according as \( \mu_0 > 1 \) or \( < 1 \). We thereby obtain the following corollary.

**Corollary 3.3b.** For the polynomial \( f(z) = a_0 + a_1 z + \cdots + a_n z^n \), let
\[
\lambda_0 = \sum_{j=0}^{n-1} \left| \frac{a_j}{a_n} \right| \quad \text{and} \quad \mu_0 = \sum_{j=1}^{n-1} \left| \frac{a_j}{a_n} \right|.
\]
Let \( \gamma = \lambda_0 \text{ or } \lambda_0^{1/n} \) according as \( \lambda_0 > 1 \) or \( < 1 \), and let \( \delta = \mu_0 \text{ or } \mu_0^{1/n} \) according as \( \mu_0 > 1 \) or \( < 1 \). Then all the zeros of \( f(z) \) lie in or on the curve \( G(\delta, \gamma; n, \alpha_0) \) where \( \alpha_0 = \arg a_0/a_n \).

4. **Generalizations.** Let us define \( \Psi_{kp}(z) = F_p(z) - |a_k| z^k, \) \( k \neq p \). Since \( \Psi_{op}(z) = z \Phi_p(z) \), the positive zeros of \( \Phi_p(z) \) are also the positive zeros of \( \Psi_{op}(z) \). By modifying somewhat the details of proof of Theorem 1.1, we may prove the following generalization.

**Theorem 4.1.** Under the hypotheses of Pellet’s Theorem the polynomial
\[
\Psi_{kp}(z) = F_p(z) - |a_k| z^k, \quad k \neq p, a_k \neq 0,
\]
has also two positive zeros \( r_k \) and \( R_k \) with \( r_k < R < R_k \). Furthermore \( f(z) \) has exactly \( p \) zeros in or on the curve \( G(r_k, r; p - k, \alpha_k) \) where \( \alpha_k = \arg (a_k/a_p) \) and none between the curves \( G(r_k, r; p - k, \alpha_k) \) and \( G(R, R_k; p - k, \alpha_k + \pi) \).
Our final generalization will consist in replacing the polynomial \( f(z) \) of Theorem 4.1 by a power series.

**Theorem 4.2.** If the power series
\[
f(z) = a_0 + a_1 z + \cdots + a_p z^p + \cdots,
\]
having a radius of convergence of \( \rho, 0 < \rho \leq \infty \), is such that each polynomial
\[
F_{np}(z) = \sum_{k=0}^{p} a_k z^k
\]
with \( n \geq N > p \) has a positive zero \( r^{(n)} \), \( r^{(n)} \leq \rho_1 < \rho \), then the function \( F_p(z) = \lim_{n \to \infty} F_{np}(z) \) has a positive zero \( r < \rho \); the function
\[
\Psi_{kp}(z) = F_p(z) - |a_k| z^k,
\]
has a positive zero \( r_k, r_k < r < \rho \), and the function \( f(z) \) has exactly \( p \) zeros in or on the curve \( G(r_k, r; p - k, \alpha_k) \) and hence in the curve \( G(r_k, \rho; p - k, \alpha_k) \).

This theorem results from Theorem 4.1 on the use of the Hurwitz theorem that within its circle of convergence a non-constant power series \( f(z) = \sum_{j=0}^{\infty} a_j z^j \) has as zeros the limit points of the zeros of the polynomials \( f_n(z) = \sum_{j=0}^{n} a_j z^j \).

If \( F_{np}(z) \) has two positive zeros in \( |z| < \rho \), we may choose \( r^{(n)} \) as the smaller one. Letting
\[
\Psi_{np}(z) = F_{np}(z) - |a_k| z^k,
\]
we see that \( \Psi_{np}(z) \) has a positive zero \( r_k^{(n)}, r_k^{(n)} < r^{(n)} \). Clearly, the power series \( F_p(z) \) and \( \Psi_{kp}(z) \) have the same radius \( \rho \) of convergence and have respectively the positive zeros \( r = \lim_{n \to \infty} r^{(n)} \) and \( r_k = \lim_{n \to \infty} r_k^{(n)} \), with \( r_k < r < \rho \). Now, given any small positive \( \epsilon \), we can find an \( N > 0 \) such that the circle of radius \( \epsilon \) drawn about the point \( z = r \) will contain \( r^{(n)} \) for all \( n \geq N \) and the circle of radius \( \epsilon \) drawn about \( z = r_k \) will contain \( r_k^{(n)} \) for all \( n \geq N \). This means that in or on the curve \( G(r_k + \epsilon, r + \epsilon; p - k, \alpha_k) \), which for any sufficiently small positive \( \epsilon \) is contained in the circle \( |z| < \rho \), lie exactly \( p \) zeros of each polynomial \( f_n(z) \) for all \( n \geq N \). Since a circle of radius \( \epsilon \) about any zero of \( f(z) \) in \( |z| < \rho \) contains a zero of each \( f_n(z), n \geq N \), it follows that in or on the curve \( G(r_k + \epsilon, r + \epsilon; p - k, \alpha_k) \) lie exactly \( p \) zeros of \( f(z) \). Since \( \epsilon \) is an arbitrary, small positive number, it follows that exactly \( p \) zeros of \( f(z) \) lie in or on the curve \( G(r_k, r; p - k, \alpha_k) \) as stated in Theorem 4.2.

**University of Wisconsin, Milwaukee**