On the Distribution of the Critical Points of a Polynomial

Jerry Shao-Chieh Cheng

June 4th 2012

Consider a polynomial $p(x)$ of a real variable x. Rolle's Theorem says that between any two distinct roots of p lies at least one root of p'. A natural question to ask is whether a similar theorem exists for a complex polynomial $p(z)$, where z is a complex variable. The answer is yes, and in fact much fascinating work has been done regarding the location of the critical points of complex polynomials. In this paper we present some of the most beautiful theorems in this subject. Much of the material here is based on Kalman's award-winning expository article *An Elementary Proof of Marden's Theorem* [3], and M. Marden's book *Geometry of Polynomials* [4].

We will first discuss the Gauss-Lucas Theorem, a generalization of Rolle's Theorem for complex polynomials. We then present Marden's Theorem, a striking special case of the Gauss-Lucas Theorem, concerning only degree three polynomials. Next, we will state Jensen's Theorem and Walsh's Two Circle Theorem, which draw surprising conclusions on the distribution of the critical points when additional constraints are imposed on the original polynomial.

1 General polynomials

We begin with some basic definitions [2, 5].

Definition 1. *A point set S is* ***convex*** *if for every two points in S, the straight line segment between the two points lies entirely in S. For a set of points $Z = \{z_1, z_2, \ldots, z_n\}$, the* ***convex hull*** *of Z, denoted H, is the smallest convex polygon that encloses Z, smallest in the sense that there is no other polygon H' such that $H \supset H' \supseteq Z$.*

Definition 2. *Let v_1, v_2 be two linearly independent vectors in $\mathbb{R}^2$. A **conical combination** of v_1, v_2 is an element of the form $\alpha_1 v_1 + \alpha_2 v_2$, where the coefficients α_i are nonnegative. The **conical hull** of v_1, v_2 is the set containing all possible conical combinations of v_1 and v_2. Geometrically, the conical hull of v_1, v_2 is the unbounded conical region in $\mathbb{R}^2$ spanned by v_1, v_2.*

We now start with the Gauss-Lucas Theorem [4, p. 22].

Theorem 1 (Gauss-Lucas). *All the critical points of a non-constant polynomial f with complex coefficients lie in the convex hull H of the set of zeros of f. If the zeros of f are not collinear, no critical point of f lies on the boundary of H unless it is a multiple zero of f.*

Before proving Theorem 1, we will first prove a general form of a result stated in Marden [4, p. 1].

Theorem 2. *If each complex number $w_j = r_j e^{i\theta_j}$, $j = 1, 2, \ldots, p$, has the property that $w_j \neq 0$, and*

$$\gamma \leq \theta_j < \gamma + \delta, \quad j = 1, 2, \ldots, p$$

where γ, δ are real constants, and $\delta < \pi$, then their sum

$$w = re^{i\theta} = \sum_{j=1}^{p} w_j$$

is nonzero. Furthermore, $\gamma \leq \theta < \gamma + \delta$ (See Figure 1).

Proof. It suffices to show the theorem for the case where $\gamma = 0$. The general result can then be obtained simply through rotation by the angle γ.

If, for all j, $0 \leq \theta_j < \delta$, then w_j lies in the conical hull of $\{1, e^{i\delta}\}$. In other words, w_j can be written as $w_j = x_j + y_j \cdot e^{i\delta}$, where $x_j, y_j \geq 0$. Let

$$w = \sum_{j=1}^{p} w_j = \sum_{j=1}^{p} x_i + (\sum_{j=1}^{p} y_i) e^{i\delta} = x + ye^{i\delta}.$$

If all the y_j's are zero, then w must lie on the positive axis, and is therefore non-zero. If there exists y_j greater than zero, then $y = \sum_{j=1}^{p} y_i > 0$, and so by the linear independence of the elements 1 and $e^{i\delta}$, $w = x + ye^{i\delta}$ is also nonzero. Furthermore, since $x, y \geq 0$, $w = re^{i\theta}$ is in the conical hull of $\{1, e^{i\delta}\}$. Therefore, $0 \leq \theta < \delta$. □

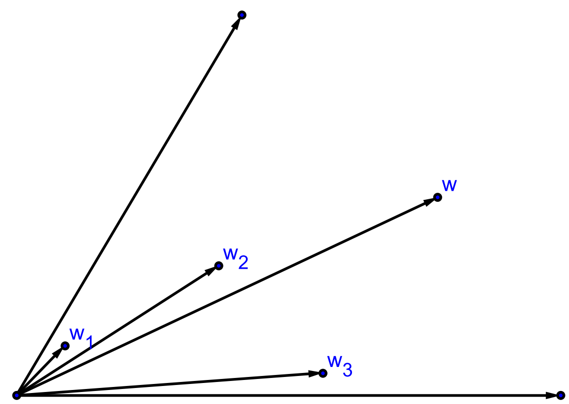

Figure 1: Illustration of Theorem 2

Armed with Theorem 2, we may now prove the Gauss-Lucas Theorem. Let

$$f(z) = (z - z_1)^{m_1}(z - z_2)^{m_2} \cdots (z - z_p)^{m_p}$$

be a polynomial with complex roots. If $f(z) \neq 0$, we may define the function $F(z)$ as the log derivative of f:

$$F(z) = \frac{f'(z)}{f(z)} = \frac{d}{dz}\log(f(z)) = \sum_{j=1}^{p} \frac{m_j}{z - z_j}.$$

Note that

$$-\overline{F(z)} = \sum_{j=1}^{p} \frac{m_j}{\bar{z}_j - \bar{z}} = \sum_{j=1}^{p} m_j w_j. \tag{1}$$

where $w_j = \dfrac{1}{\bar{z}_j - \bar{z}}$.

If $z_j - z = r_j e^{i\theta_j}$, then a brief calculation shows that

$$w_j = \frac{1}{\bar{z}_j - \bar{z}} = \frac{1}{r_j} e^{i\theta_j}$$

Thus, if we let $w_j = \rho_j e^{i\phi_j}$, then $\phi_j = \theta_j$ and $\rho = 1/r_j$.

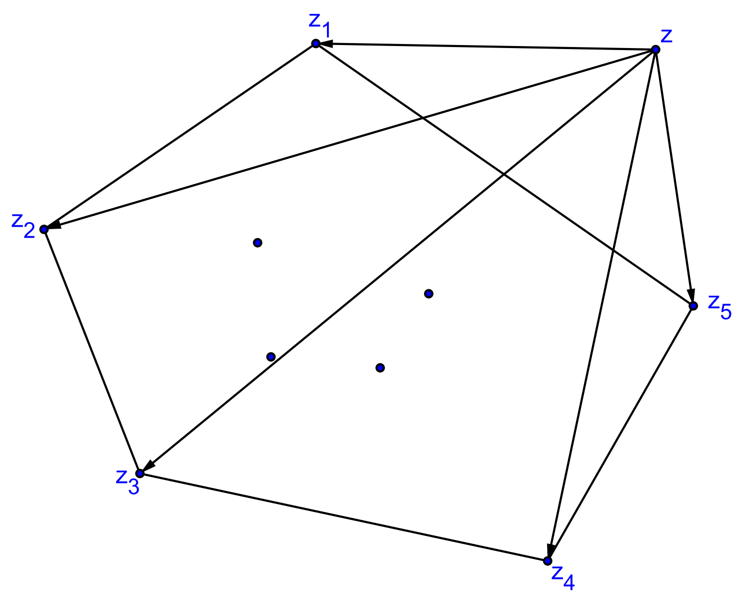

Figure 2: Illustration of the Gauss-Lucas Theorem

Now suppose, for contradiction, that f' has a zero $z \notin H$. Since H is a convex hull, the vectors $z_j - z = r_j e^{i\theta_j}$ satisfy

$$\gamma \leq \theta_j < \gamma + \delta$$

for all j, where γ, δ are real constants, and $\delta < \pi$ (see Figure 2). This implies that for each $w_j = \rho_j e^{i\phi_j}$, we have $\gamma \leq \phi_j < \gamma + \delta$. Thus, if we write $-\overline{F(z)}$ in terms of Equation 1, the w_j's will satisfy the hypothesis in Theorem 2. Hence $-\overline{F(z)} \neq 0$, and so

$$F(z) = \frac{f'(z)}{f(z)} \neq 0$$

Therefore $f'(z) \neq 0$, which is a contradiction. Hence the first part of the theorem is established.

To prove the second part of the theorem, suppose the zeros of f are not collinear. Then H must be a convex polygon. Let us name the vertices of H by $z_1, z_2, \ldots, z_p$ in a counterclockwise fashion. If z is a point lying between two vertices on an edge of H, say $\overline{z_1 z_p}$, then for each number $z_j - z = r_j e^{i\theta_j}$, there must exist γ', δ' such that $\theta_j \in [\gamma', \gamma' + \delta')$ for $j = 1, 2, \ldots, p-1$, and $\theta_p \notin [\gamma', \gamma' + \delta')$.

Let $S_{p-1} = \sum_{j=1}^{p-1} m_j w_j = \rho e^{i\phi}$. By Theorem 2, we have $S_{p-1} \neq 0$, and $\phi \in [\gamma', \gamma + \delta')$. Furthermore, since $w_p \neq 0$, and $\phi_p \notin [\gamma', \gamma + \delta')$, we know that S_{p-1} and w_p are linearly independent. Hence we have

$$-\overline{F(z)} = \sum_{j=1}^{p} m_j w_j = S_{p-1} + m_p w_p \neq 0.$$

Therefore, no critical point of f may lie on the boundary of H unless it is a vertex of H.

This completes the proof of Gauss-Lucas Theorem.

2 Marden's Theorem

We now state Marden's Theorem, a beautiful special case of the Gauss-Lucas theorem [3].

Theorem 3 (Marden)**.** *Let $p(z)$ be a third-degree polynomial with complex coefficients, and whose roots z_1, z_2, and z_3 are noncollinear points in the complex plane. Let T be the triangle with vertices at z_1, z_2, and z_3. There is an unique ellipse inscribed in T and tangent to the sides at their midpoints. The foci of this ellipse are the roots of $p'(z)$ (See Figure 3).*

An elegant proof was given in Kalman's paper [3]. The proof only involves several elementary geometric properties of ellipses and some simple analysis.

Basically, the argument can be broken down into two parts: First, we show that if an ellipse E has the roots z_1', z_2' of $p'(z)$ as its foci, and if E passes through the midpoint of one side of T, then E is actually tangent to that side. Then, we show that E is actually also tangent to the other two sides of T at their respective midpoints. Marden's Theorem will then be established.

Before we start the proof of Marden's Theorem, we present an important proposition that will greatly simplify our analysis [3].

Proposition 1. *Let $M(z) = \alpha z + \beta$ be a linear transformation in the complex plane, where $\alpha \neq 0, \beta$ are arbitrary complex numbers. Let $p(z) = (z - z_1)(z - z_2)(z - z_3)$. Let z_1', z_2' be the two roots of $p'(z)$. If we define a new polynomial $p_M(z) = (z - M(z_1))(z - M(z_2))(z - M(z_3))$, then the roots of $p_M'(z)$ are equal to $M(z_1')$ and $M(z_2')$.*

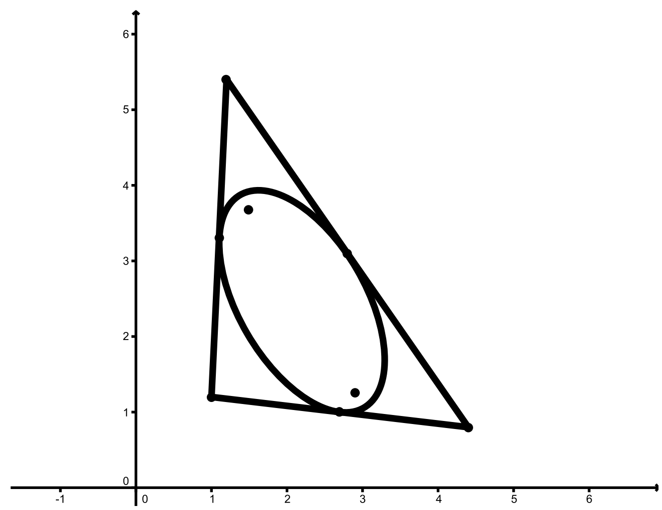

Figure 3: Illustration of Marden's Theorem

Proof. The composition function $p_M(M(z))$ can be written as

$$\begin{aligned} p_M(M(z)) &= (M(z) - M(z_1))(M(z) - M(z_2))(M(z) - M(z_3)) \\ &= \alpha^3(z - z_1)(z - z_2)(z - z_3) \\ &= \alpha^3 p(z). \end{aligned}$$

Thus, differentiating both sides with respect to z yields

$$\alpha p'_M(M(z)) = \alpha^3 p'(z)$$

So $p'_M(M(z'_1)) = p'_M(M(z'_2)) = 0$. $\square$

In other words, Proposition 1 tells us that if the roots of p are transformed under a linear map M, then the geometric configuration of the roots of $p'_M(z)$ relative to the roots of p_M stays invariant. Therefore, in the context of Marden's Theorem, we may choose a suitable linear map M, so that the roots of p are scaled, rotated, and translated by M to more convenient positions in the $\mathbb{C}$ plane. Then, it will suffice to prove the theorem with the simpler polynomial p_M.

We next demonstrate several crucial properties related to the ellipses.

Proposition 2. *Let F_1, F_2 be two distinct points on the Cartesian plane. Let L be an arbitrary straight line that does not intersect the line segment $\overline{F_1F_2}$.*

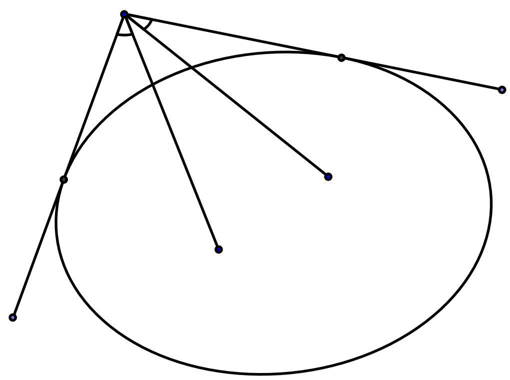

Figure 4: Illustration of Theorem 5

Then there is one and only one ellipse E with foci F_1, F_2 that is tangent to L.

Proof. See Section A at the end of the paper. □

The next two theorems are known as the "optical" properties of the ellipses. For the proofs of these theorems, see [1, p. 8] and [3].

Theorem 4. *Consider an ellipse with foci F_1 and F_2, and suppose the line L is tangent to the ellipse at point P. Then the line segments $\overline{F_1P}$ and $\overline{F_2P}$ make equal acute angles with the line L.*

In other words, if the boundary of the ellipse is a reflexive surface, then any ray issuing from one focus will be reflected by the boundary and then pass through the other focus. Theorem 4 is equivalent to the following theorem [3].

Theorem 5. *Consider an ellipse with foci F_1 and F_2, and a point A outside the ellipse. There are two lines through A that are tangent to the ellipse. Let G_1 and G_2 be the points of tangency of these lines with the ellipse. Then $\angle F_1AG_1 = \angle F_2AG_2$ (See Figure 4).*

We are now ready to prove Theorem 3. We introduce our first lemma [3].

Lemma 1. *Let the polynomial $p(z)$, its roots z_1, z_2, z_3, and the triangle T be as in the statement of Marden's Theorem. Then the ellipse E with foci at the roots of p' and passing through the midpoint of one side of the triangle T is actually tangent to that side of T.*

Proof. By Proposition 1, we can transform the roots of the polynomial by a linear map without compromising their geometric relations with the critical points. Therefore, we may choose to work with the polynomial $p(z) = (z + 1)(z - 1)(z - w)$, where $w = re^{i\theta}$, $0 < \theta < \pi$. Let z_1', z_2' be the roots of $p'(z)$, and let E be the ellipse with foci z_1', z_2' that passes through the origin O (See Figure 5). Our goal is to show that the real axis is tangent to E at O.

The derivative p' is given by

$$p'(z) = 3z^2 - 2wz - 1$$

By the Gauss-Lucas Theorem, we know that z_1', z_2' both lie inside T. Furthermore, by examining the coefficients of p', we know that $z_1'z_2' = -1/3$. Thus, if $z_1' = r_1e^{i\theta_1}$, $z_2' = r_2e^{i\theta_2}$, then we have

$$\theta_1 + \theta_2 = \pi$$

In other words, θ_1 and θ_2 are complementary angles, and so the two line segments $\overline{Oz_1'}$, $\overline{Oz_2'}$ make equal acute angles with the real axis. By Theorem 4, the real axis coincide with the line tangent to the ellipse E at the origin O. This completes the proof of the lemma. □

The second part of the proof of Marden's Theorem aims to show that the ellipse E as defined in Lemma 1 is actually tangent to the other two sides of T *at their respective midpoints* as well. This brings us to our second lemma [3]:

Lemma 2. *Let the polynomial $p(z)$, its roots z_1, z_2, z_3, and the triangle T be as in the statement of Marden's Theorem. Consider the ellipse with foci at the roots of p and which is tangent to one side of T at its midpoint. Then this same ellipse is tangent to the other two sides of T at their respective midpoints.*

Proof. This time we choose to work with the polynomial $p(z) = z(z-1)(z - w)$, where w is an arbitrary complex number in the upper half plane. Let z_4, z_5 be the roots of $p'(z)$, and let E be the ellipse with foci z_4, z_5 that is tangent to T at the point $z = 1/2$. We will first show that E is also tangent to the side Ow at the point $z = w/2$. The derivative p' is given by

$$p'(z) = 3z^2 - 2(1 + w)z + w$$

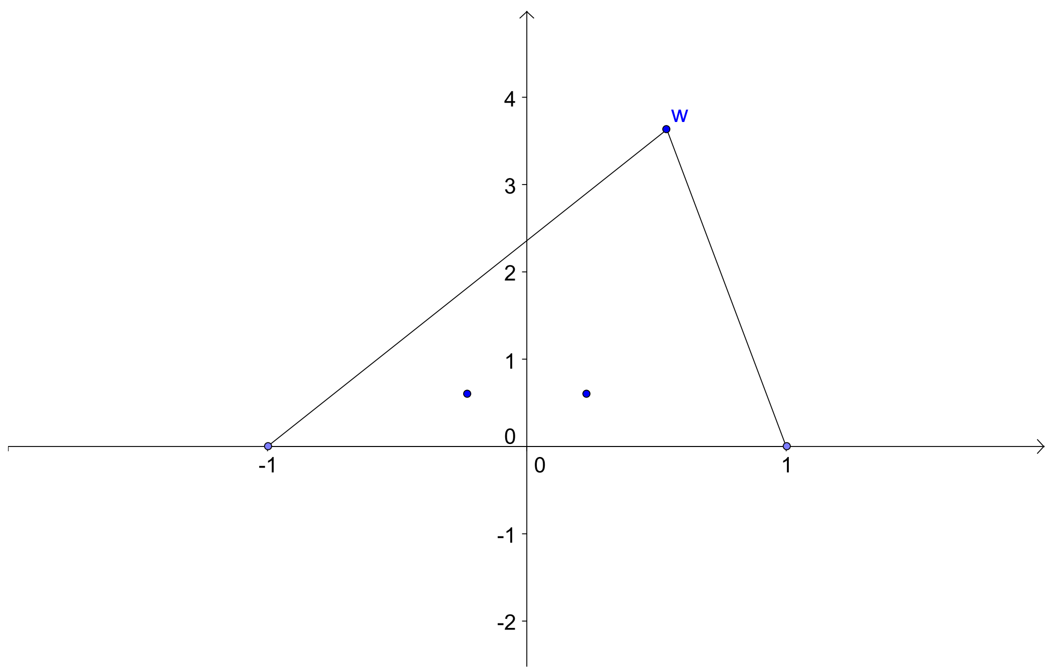

Figure 5: Illustration of Lemma 1

Again, by the Gauss-Lucas Theorem, we know that z_4, z_5 both lie inside T. Furthermore, by examining the coefficients of p', we know that $z_4 z_5 = -w/3$. Thus, if $z_4 = r_4 e^{i\theta_4}$, $z_5 = r_5 e^{i\theta_5}$, where $0 < \theta_4 \leq \theta_5 < \pi$, we have

$$\theta_4 + \theta_5 = \operatorname{Arg}(w)$$

Hence,

$$\begin{aligned} \angle wOz_5 &= \operatorname{Arg}(w) - \operatorname{Arg}(z_5) \\ &= (\theta_4 + \theta_5) - \theta_5 \\ &= \theta_4 \\ &= \angle z_4 OB \end{aligned}$$

Since E is tangent to the real axis at the point $z = 1/2$, we know that O lies outside E. Now suppose the ray Ow' is the other tangent line to E. Then, by Theorem 5, $\angle w'Oz_5 = \angle z_4 OB$. But we have just shown that $\angle z_4 OB = \angle wOz_5$, so $\angle w'Oz_5 = \angle wOz_5$. This means that the line Ow coincides with the tangent line Ow'. So E is tangent to Ow.

To see that the point of tangency of Ow on E is the midpoint of Ow, let E' be the ellipse with foci z_4, z_5 that is tangent to Ow at $z = w/2$. By Lemma 1, E' is also tangent to the real axis. Thus, by Proposition 2, E' and E are in fact identical, and so E is tangent to Ow at its midpoint.

That E is also tangent to the side Bw at its midpoint, where B is the point $z = 1$, can be established similarly by working with another polynomial $q(z) = z(z+1)(z-w')$, where w' lies in the upper half plane.

Thus the lemma is established, and the proof of Marden's theorem is complete. □

3 Polynomials with Real Coefficients

Both the Gauss-Lucas Theorem and Marden's Theorem deal with a polynomial p with complex coefficients. What happens if we restrict the coefficients of p to be real? The first thing to notice is that the roots of $p(z)$ come in conjugate pairs. Does this additional constraint on the roots allow us to draw a better conclusion on the distribution of the critical points? The answer is yes, as seen in Jensen's Theorem [4, p. 25].

Let $p(z)$ be a polynomial of a complex variable with real coefficients. For each non-real conjugate pair $x_j \pm iy_j$ of the roots of p, we construct a disk centered at the point $z = x_j$ with radius y_j. The disks thus formed are called the *Jensen disks* of p (See Figure 6).

Theorem 6 (Jensen)**.** *Every non-real critical point of p must lie in one of the Jensen disks of p.*

Proof. Let $z = x + iy$ be a non-real complex number lying outside all of the Jensen disks of p. It is clear that $p(z) \neq 0$. We will show that $p'(z) \neq 0$.

Suppose

$$p(z) = (z - z_1)(z - z_2) \cdots (z - z_n)$$

where each zero of multiplicity m is listed m times. The log derivative of p is given by

$$F(z) = p'(z)/p(z) = \sum_{j=1}^{n} \frac{1}{z - z_j} = \sum_{Im(z_j)=0} \frac{1}{z - z_j} + \sum_{Im(z_j)\neq 0} \frac{1}{z - z_j}$$

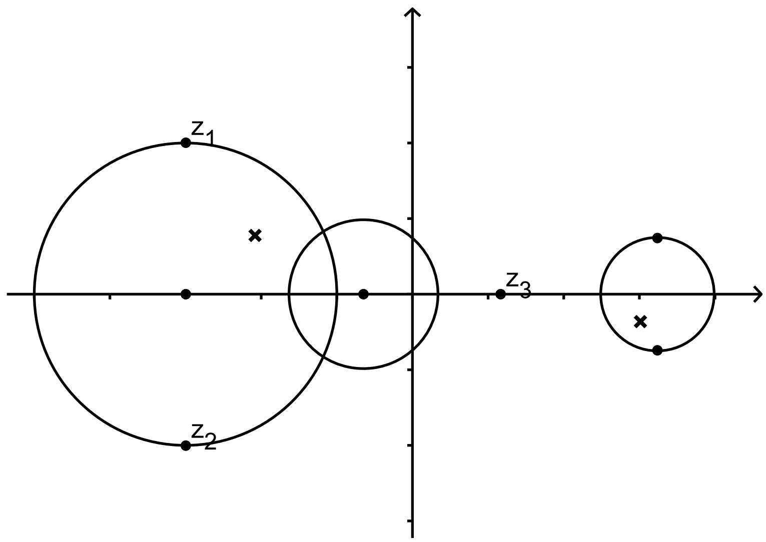

Figure 6: Illustration of Theorem 6

For each j, let $w_j = \dfrac{1}{z - z_j}$. Now suppose p has a conjugate pair of roots, which we may call z_1 and z_2. Let $z_1 = x_1 + iy_1$, $z_2 = x_1 - iy_1$. Then, after some calculation, we get

$$Im(w_1 + w_2) = \frac{-2y[(x - x_1)^2 + y^2 - y_1^2]}{[(x - x_1)^2 + (y - y_1)^2][(x - x_1)^2 + (y + y_1)^2]}$$

Since $z = x + iy$ lies outside the Jensen disk centered at x_1 with radius y_1, we have

$$(x - x_1)^2 + y^2 > y_1^2$$

Therefore we obtain

$$\operatorname{sgn}(Im(w_1 + w_2)) = -\operatorname{sgn}(z)$$

On the other hand, suppose p has a real root, which we may call $z_3 = x_3 + 0 \cdot i$. Then, after some calculation, we get

$$Im(w_3) = \frac{-y}{(x - x_3)^2 + y^2}.$$

which means that

$$\operatorname{sgn}(Im(w_3)) = -\operatorname{sgn}(z)$$

Therefore, the sign of the imaginary part of $F(z) = \sum_{j=1}^{n} w_j$ is always opposite to the sign of the imaginary part of z, which is nonzero if z is not a real number. So we can conclude that $F(z)$ is nonzero, and therefore $p'(z)$ is nonzero. This completes the proof. $\square$

4 Walsh's Two Circle Theorem

We now return our attention to a complex polynomial p, but this time imposing another kind of restriction on its roots. More specifically, by the Gauss-Lucas Theorem, we know that if all the roots of p can be enclosed in a circle, then its critical points stay inside the same circle as well. But what happens if the roots of p form two separate clusters, and we can use two circles to enclose all the roots of p? Walsh's Two Circle Theorem gives us an surprising answer to this question.

Before presenting the theorem, we introduce several important facts about polynomials and linear functions.

Let $f(z) = \sum_{\nu=0}^{n} a_\nu z^\nu$, $g = \sum_{\nu=0}^{n} b_\nu z^\nu$ be two polynomials. We make a slight change of notation and write

$$a_k = \binom{n}{k} A_k, \quad b_k = \binom{n}{k} B_k$$

for $k = 0, 1, \ldots, n$, where A_k, B_k are suitable constants.

We now provide a definition of a relationship between two polynomials of the same degree [4, p. 60]:

Definition 3. *Two polynomials*

$$f(z) = \sum_{k=0}^{n} \binom{n}{k} A_k z^k, \quad g(z) = \sum_{k=0}^{n} \binom{n}{k} B_k z^k, \quad A_n B_n \neq 0$$

are said to be ***apolar*** *if their coefficients satisfy the equation*

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} A_k B_{n-k} = 0.$$

Example 1. *The polynomials* $f(z) = 3z^3 - 18z^2 + 42z - 36$ *and* $g(z) = z^3 - 6z^2 + 11z - 6$ *are apolar polynomials.*

The following theorem, due to Grace, describes the relationship between the roots of two apolar polynomials. For a proof, see [4, p. 61].

Theorem 7 (Grace)**.** *If $f(z)$ and $g(z)$ are apolar polynomials, and if one of them has all its zeros in a cicular region C, then the other will have at least one zero in C.*

Next, we present some fundamental notions regarding symmetric linear forms [4, p. 60], [7, p. 181].

Definition 4. *The* ***elementary symmetric polynomials*** *in the variables $z_1, z_2, \ldots, z_n$, denoted $s(n,p)$, are defined by*

$$\begin{aligned} s(n,0) &= 1, \\ s(n,p) &= \sum_{1 \le j_1 < j_2 < \cdots < j_p \le n} z_{j_1} z_{j_2} \cdots z_{j_p}, \quad 1 \le p \le n. \end{aligned}$$

For a polynomial $f(z) = A_n \prod_{j=1}^{n}(z - z_j) = \sum_0^n \binom{n}{k} A_k z^k$, the coefficients are related to the elementary symmetric polynomials of its zeros in the following way:

$$\binom{n}{p} A_{n-p} = (-1)^p s(n,p) A_n, \quad 0 \le p \le n. \tag{2}$$

Therefore, we can write the coefficients of f in terms of the elementary symmetric polynomials:

$$f(z) = A_n \sum_{p=0}^{n} [(-1)^{n-p} s(n, n-p)] z^p \tag{3}$$

Definition 5. *A* ***symmetric*** n***-form*** *is a polynomial $L(z_1, z_2, \ldots, z_n)$ in the complex variables $z_1, z_2, \ldots, z_n$ which is linear in each of its variables, i.e. if the variables $z_k (k \neq j)$ are kept fixed, the resulting polynomial in z_j is of first degree. The form is* ***symmetric*** *if L is independent of any permuation of the n variables.*

Example 2. *The elementary symmetric polynomials in the variables $z_1, z_2, \ldots, z_n$ are symmetric linear n-forms. Hence, by Equation 3, the polynomial $f(z) = (z - z_1)(z - z_2) \cdots (z - z_{n_1})$ is a symmetric linear n-form in the variables $z_1, z_2, \ldots, z_n$. Differentiating f term by term, we may also see that f' is a symmetric linear n-form.*

The next theorem is a special case of the Fundamental Theorem of Symmetric Polynomials. For a proof, see [6, p. 7]

Theorem 8. *For every symmetric linear n-form $L(z_1, z_2, \ldots, z_n)$ there exists constants B_k, $0 \le k \le n$, such that*

$$L(z_1, z_2, \ldots, z_n) = \sum_{k=0}^{n} B_k \cdot s(n,k)$$

where $s(n,k)$ are the elementary symmetric polynomials in the variables $z_1, z_2, \ldots, z_n$.

Combining Theorem 7 and 8, we may prove the Walsh Coincidence Theorem [4, p. 62]:

Theorem 9 (Walsh Coincidence Theorem)**.** *Let Φ be a symmetric linear n-form in $z_1, z_2, \ldots z_n$ and let C be a circular region containing the n points $z_1^{(0)}, z_2^{(0)}, \cdots, z_n^{(0)}$. Then in C there exists at least one point ζ such that*

$$\Phi(\zeta, \zeta, \ldots, \zeta) = \Phi(z_1^{(0)}, z_2^{(0)}, \cdots, z_n^{(0)}).$$

Proof. Let $\Phi(z_1^{(0)}, z_2^{(0)}, \cdots, z_n^{(0)}) = \Phi_0$. Then $\Phi(z_1, z_2, \ldots, z_n) - \Phi_0$ is a symmetric linear n-form in $z_1, z_2, \ldots, z_n$. By Theorem 8, there exist constants B_k such that

$$\begin{aligned}\Phi(z_1, z_2, \ldots, z_n) - \Phi_0 &= \sum_{k=0}^{n} B_k \cdot s(n,k) \\ &= \frac{1}{A_n} \sum_{k=0}^{n} (-1)^k \binom{n}{k} A_{n-k} B_k\end{aligned}$$

where $f(z) = A_n \prod_{j=1}^{n} (z - z_j) = \sum_0^n \binom{n}{k} A_k z^k$. The second equality above is obtained by referring to Equation 2 in Definition 4. In particular, since

$$\Phi(z_1^{(0)}, z_2^{(0)}, \cdots, z_n^{(0)}) - \Phi_0 = 0,$$

the polynomial $f(z) = A_n \prod_{j=1}^{n} (z - z_j^{(0)})$ is apolar to the polynomial

$$g(z) = \sum_{k=1}^{n} \binom{n}{k} B_k z^k = \Phi(z, z, \ldots, z) - \Phi_0.$$

By Theorem 7, $g(z) = \Phi(z, z, \ldots, z) - \Phi_0$ has a root ζ in the circular region C. This completes the proof. $\square$

Next, we present a lemma that is crucial for working with the statement of Walsh's Two Circle Theorem [4, p. 75].

Lemma 3. *Let C_1, C_2 be circles of radius r_1, r_2 centered at c_1, c_2, respectively. If $\alpha_1 \in C_1$, $\alpha_2 \in C_2$, then for all positive real numbers m_1, m_2, the point $\alpha = m_1\alpha_1 + m_2\alpha_2$ lies in the circle C with radius $r = m_1r_1 + m_2r_2$ centered at $c = m_1c_1 + m_2c_2$.*

Proof. Left as an exercise to the reader.

□

Finally, we state and prove the Continuity Theorem, which says that the zeros of a polynomial vary continuously with its coefficients [6, p. 9]. As we will see presently, this theorem furnishes an elegant proof to Walsh's Two Circle Theorem.

Theorem 10 (Continuity Theorem)**.** *Let*

$$f(z) = \sum_{\nu=0}^{n} a_v z^\nu = \prod_{j=1}^{k}(z - z_j)^{m_j} \quad (m_1 + \cdots + m_k = n)$$

be a monic polynomial of degree n with distinct zeros $z_1, \ldots, z_k$ of multiplicities $m_1, \ldots, m_k$. Then, given a positive $\epsilon < \min_{1<i<j\le k} |z_i - z_j|/2$, there exists a $\delta > 0$ so that any monic polynomial $g(z) = \sum_{\nu=0}^{n} b_\nu z^\nu$ whose coefficients satisfy $|b_\nu - a_\nu| < \delta$, for $\nu = 1, \ldots, n-1$, has exactly m_j zeros in the disk

$$D_j = \{z : |z - z_j| < \epsilon\} \quad (j = 1, \ldots, k).$$

Proof. Fix an ϵ smaller than $\min_{1<i<j\le k} |z_i - z_j|/2$. Since the zeros of f are isolated, we may pick $0 < \rho < \epsilon$ such that $|f(\zeta)| \neq 0$ for all ζ satisfying $|\zeta - z_j| \le \rho$.

Now we seek a δ such that if $h(z) = \sum_{\nu=0}^{n} c_\nu z^\nu$, where $|c_\nu| < \delta$ for all ν, then $f + h$ and f have the same number of zeros in the disk $D(z_j; \rho)$ for all j.

We find δ as follows. Let U be the union of the circles $C_j = \{z : |\zeta - z_j| = \rho\}$. Since U is a compact set, we may find real positive constants M, m such that $\sum_{\nu=0}^{n} |\zeta|^n \le M$ and $|f(\xi)| \ge m$ for all $\zeta, \xi \in U$. If we choose δ small so that $\delta M < m$, then

$$|h(\zeta)| \le \sum_{\nu=0}^{n} |c_\nu||z^\nu| < \delta M < m \le |f(\zeta)|$$

for all ζ on each circle C_j. By Rouche's Theorem, $f+h$ and f have the same number of zeros in D_j for all j, and our result follows. $\square$

At last, we are ready to state and prove Walsh's Two Circle Theorem. This beautiful theorem as presented here is taken from [7, p. 188]. The proof is from [4, p. 89].

Theorem 11 (Walsh's Two Circle Theorem)**.** *Let C_1 be a disk with center c_1 and radius r_1 and C_2 a disk with center c_2 and radius r_2. Let P be a polynomial of degree n with all its zeros in $C_1 \cup C_2$, say n_1 zeros in C_1 and n_2 zeros in C_2. Then P has all its critical points in $C_1 \cup C_2 \cup C_3$, where C_3 is the disk with center c_3 and radius r_3 given by*

$$c_3 = \frac{n_1 c_2 + n_2 c_1}{n}, \quad r_3 = \frac{n_1 r_2 + n_2 r_1}{n}$$

Furthermore, if C_1, C_2, and C_3 are pairwise disjoint, then C_1 contains $n_1 - 1$ critical points, C_2 contains $n_2 - 1$ critical points, and C_3 contains 1 critical point (See Figure 7).

Proof. Let

$$f(z) = (z - z_1)(z - z_2)\cdots(z - z_{n_1})(z - \xi_1)(z - \xi_2)\cdots(z - \xi_{n_2})$$

where the z_i's are the zeros of f in C_1, and the ξ_i's are the zeros of f in C_2. Each zero of multiplicity m is listed m times. Let $f = f_1 f_2$ where $f_1(z) = \prod(z - z_i)$ and $f_2(z) = \prod(z - \xi_i)$. Then if Z is a critical point of f, we have

$$0 = f'(Z) = f_1'(Z) f_2(Z) + f_1(Z) f_2'(Z) \tag{4}$$

By the remarks in Example 2, this equation is a symmetric linear n-form in the zeros $z_1, z_2, \ldots z_{n_1}, \xi_1, \xi_2, \ldots, \xi_{n_2}$. By Walsh's Coincidence theorem, there exists $\zeta_1 \in C_1$, $\zeta_2 \in C_2$ such that in Equation 4, each of the variables $z_1, z_2, \ldots z_{n_1}$ maybe replaced by ζ_1 , and similarly each of the variables $\xi_1, \xi_2, \ldots, \xi_{n_2}$ maybe replaced by ζ_2. In other words, we have

$$\begin{aligned} 0 &= n_1(Z - \zeta_1)^{n_1 - 1}(Z - \zeta_2)^{n_2} + n_2(Z - \zeta_1)^{n_1}(Z - \zeta_2)^{n_2 - 1} \\ &= (Z - \zeta_1)^{n_1 - 1}(Z - \zeta_2)^{n_2 - 1}[n_1(Z - \zeta_2) + n_2(Z - \zeta_1)] \end{aligned}$$

Thus, the only possible values of Z are

$$Z = \zeta_1 \in C_1 \quad (\text{if } n_1 > 0), \quad Z = \zeta_2 \in C_2 \quad (\text{if } n_2 > 0),$$

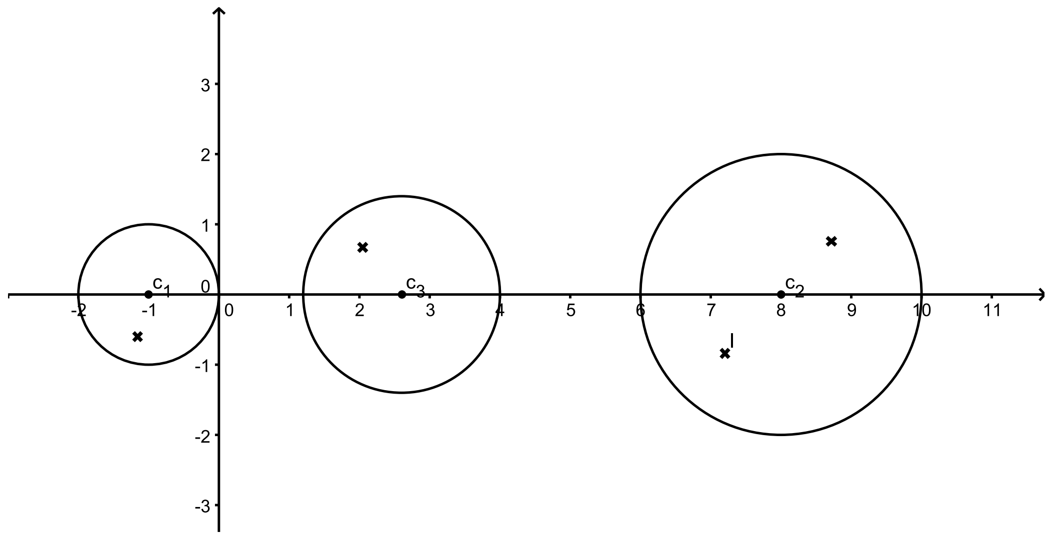

Figure 7: Illustration of Walsh's Two Circle Theorem

or

$$Z = \frac{n_1\zeta_2 + n_2\zeta_1}{n_1 + n_2}.$$

By setting $m_1 = \dfrac{n_2}{n_1 + n_2}$, $m_2 = \dfrac{n_1}{n_1 + n_2}$ and applying Lemma 3, we see that the point $Z = \dfrac{n_1\zeta_2 + n_2\zeta_1}{n_1 + n_2}$ must be in C_3.

Therefore every critical point of f is in $C_1 \cup C_2 \cup C_3$. This establishes the first part of the theorem.

The proof for the second part of the result is especially interesting. Assume C_1, C_2, and C_3 are disjoint. Our goal is to show that C_1 contains $n_1 - 1$ critical points, C_2 contains $n_2 - 1$ critical points, and C_3 contains 1 critical point. To this end, we examine the zeros of f' as we cluster the zeros of f in the following way: First we move $z_1, z_2, \ldots z_{n_1}$ continuously to a single point $t_1 \in C_1$, and then we move $\xi_1, \xi_2, \ldots, \xi_{n_2}$ continuously to a single point $t_2 \in C_2$. Then f becomes the polynomial $(z - t_1)^{n_1}(z - t_2)^{n_2}$, and so f' has zero at t_1 with multiplicity $n_1 - 1$, and at t_2 with multiplicity $n_2 - 1$. This means that f' (which has degree $n_1 + n_2 - 1$) has exactly one zero in C_3.

However, by Equation 3 in Definition 4, the coefficients of f' can be expressed in terms of the elementary symmetric polynomials in the variables $z_1, z_2, \ldots z_{n_1}, \xi_1, \xi_2, \ldots, \xi_{n_2}$. Hence, by Theorem 10, the zeros of f' depend continuously on the zeros of f. This means that during the clustering process, no zero of f' may "jump" from one circle to another circle. In other words, the net change in the number of zeros of f' in each of the mutually disjoint circles C_1, C_2 and C_3 is 0. Therefore, the derivative of the polynomial we started with (before the clustering process) must also have exactly $n_1 - 1$ zeros in C_1, $n_2 - 1$ zeros in C_2, and one zero in C_3. This completes the proof for the second half of the theorem. $\square$

5 Conclusion

We have presented and proved four important theorems in the study of critical points of polynomials. The Gauss-Lucas Theorem is of fundamental importance, and is used multiple times throughout the paper. Meanwhile, the proof of Marden's Theorem utilizes many remarkable tools. In particular, the idea of re-positioning the roots of a polynomial greatly simplifies our analysis. Next, Jensen's Theorem presents an interesting improvement of the Gauss-Lucas Theorem when the polynomials have only real coefficients. Finally, we prove the beautiful Walsh's Two Circle Theorem, which involves an elegant use of the continuity of critical points.

We hope that the theorems and techniques discussed in this paper are not only asthetically pleasing, but also potentially adoptable by the reader to their future mathematical endeavors.

References

[1] A. V. Akopyan, A. A. Zaslavskiĭ, *Geometry of Conics.* American Mathematical Soc., Dec 13, 2007.

[2] J. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms.* Springer, Oct 21, 1993.

[3] D. Kalman, *An Elementary Proof of Marden's Theorem.* The American Mathematical Monthly, vol. 115, no. 4, April 2008, pp. 330-338

[4] M. Marden, *Geometry of Polynomials.* Math. Surveys no. 3, American Mathematical Society, Providence, RI, 1966.

[5] J. O'Rourke, *Computational Geometry in C.* Cambridge University Press, Oct 13, 1998.

[6] Q. I. Rahman, G. Schmeisser, *Analytic Theory of Polynomials.* Clarendon Press, Nov 7, 2002.

[7] T. Sheil-Small, *Complex Polynomials.* Cambridge University Press, Nov 7, 2002.

A Appendix: Proof of Proposition 2

We prove Proposition 2 in this section:

Proposition. *Let F_1, F_2 be two distinct points on the Cartesian plane. Let L be an arbitrary straight line that does not intersect the line segment $\overline{F_1F_2}$. Then there is one and only one ellipse E with foci F_1, F_2 that is tangent to L.*

Proof. Without loss of generality, we may assume that $F_1 = (1,0)$, $F_2 = (-1,0)$. Let $L : y = mx + b$. The constraint that L does not intersect $\overline{F_1F_2}$ is equivalent to the relation $b^2 > m^2$. If E is an ellipse with foci F_1, F_2, it can be described by the equation

$$E : \frac{x^2}{a^2} + \frac{y^2}{a^2 - 1} = 1, \quad a > 1$$

Substituting $y = mx + b$ into the equation describing E yields a quadratic equation in x:

$$(a^2 - 1 + a^2m^2)x^2 + 2ma^2bx + [a^2b^2 - a^2(a^2 - 1)] = 0$$

If the line L is tangent to E, the quadratic above must have a double root. This happens if and only if the discriminant of the quadratic is 0:

$$\begin{aligned}\Delta_x &= (2ma^2b)^2 - 4(a^2 - 1 + a^2m^2)[a^2b^2 - a^2(a^2 - 1)] \\ &= 4a^2[(m^2 + 1)a^4 + (-2 - m^2 - b^2)a^2 + (b^2 + 1)] \\ &= 0\end{aligned}$$

We must show that there is a unique parameter $a > 1$ such that $\Delta_x = 0$. Let $\tau = a^2$. Then $\Delta_x = 0$ iff

$$(m^2 + 1)\tau^2 + (-2 - m^2 - b^2)\tau + (b^2 + 1) = 0$$

Treated as a quadratic in τ, this equation has an obvious solution $\tau_1 = 1$, which yields $a = \pm 1$. But ± 1 are not valid candidates for a. On the other hand, since we know the sum of the two roots $\tau_1 + \tau_2 = -(-2 - m^2 - b^2)/(m^2 + 1) = 1 + (1 + b^2)/(1 + m^2)$, the value of the other root τ_2 is $(1 + b^2)/(1 + m^2)$. The constraint $b^2 > m^2$ implies that $\tau_2 > 1$. So another two choices for a are $a = \pm\sqrt{\tau_2}$, and only the positive value is a valid choice for a. Thus, we have shown that there is a unique $a > 1$ such that $\Delta_x = 0$.

Hence, there is only one ellipse E with foci F_1, F_2 that is tangent to L. $\square$