Critical Points of Complex Polynomials

Taboka Prince CHALEBGWA (taboka@aims.ac.za)

African Institute for Mathematical Sciences (AIMS)

Supervised by: Professor Florian Breuer
Stellenbosch University, South Africa

22 May 2014

Submitted in partial fulfillment of a structured masters degree at AIMS South Africa

Abstract

The relationship between the roots and the critical points of a polynomial has been studied extensively for a considerable amount of time, consequently, many results have been achieved in this regard. In this essay, we prove and give a refinement of a result of Gauss and Lucas, which describes the location of the critical points in terms of the polygon whose vertices are the roots of the complex polynomial $p(z)$. We then study Marden's theorem, which is a special case of the Gauss-Lucas theorem, relating the roots and the critical points of cubic polynomials in terms of a geometric configuration of a triangle and an inscribed ellipse. Motivated by the results, we finally discuss Sendov's conjecture, which is a further attempt to accurately describe this relationship, we give a proof of a special case of the conjecture, and suggest a possible method towards resolving the general conjecture.

Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Chalebgwa T. P

Taboka Prince Chalebgwa, 22 May 2014

Contents

1. Introduction

Rolle's Theorem says that between any two roots of a real valued polynomial $p(x)$ there lies a critical point c of $p(x)$. The celebrated Gauss-Lucas Theorem (Curgus and Mascioni, 2004), which states that the critical points of a polynomial $p(z)$ lie in the convex hull of its roots, is a generalisation of Rolle's theorem to complex polynomials. Consequent to the Gauss-Lucas theorem, many attempts have been made to obtain increasingly accurate results in an effort to pinpoint the exact location of the critical points of a polynomial in relation to its roots.

Perhaps the most significant of these attempts is Sendov's Conjecture, posed by Blagovest Sendov in 1958, and first appeared in (Hayman, 1967). It states that if all the roots of $p(z)$ lie in a closed unit disc, then within a unit distance of each root lies at least one critical point of $p(z)$.

Marden's Theorem[1], described in (Kalman, 2008) as "the most marvellous theorem in mathematics", gives the geometrical relationship between the roots of a cubic polynomial with complex coefficients and its critical points. By the Gauss-Lucas theorem, we know that the critical points lie in the triangle formed by the roots, Marden's theorem, however, precisely identifies these critical points as the foci of a unique ellipse inscribed within the triangle and tangent to the three sides.

(Frayer et al., 2014) considered this special case of cubic polynomials, further exploring the structure of the critical points, by studying the behaviour of the critical points of a polynomial as its roots are dragged. The authors obtain an existence and uniqueness characterisation of a certain class of cubic polynomials in terms of the locations of the roots and critical points relative to each other. This less common approach starts by defining locations of the critical points, and then determines the possible classes of polynomials corresponding to them. This is the converse of the more common approach of starting with a polynomial, then studying the relative locations of its roots and critical points.

On the other hand, in (Curgus and Mascioni, 2004), the authors obtained an improvement of the Gauss-Lucas theorem, showing that there is an area in the convex hull of the roots of $p(z)$ which can be excluded in the search of the location of the critical points, this gives new predictions for regions in which we are guaranteed to find the critical points.

In this essay we begin with a study of the results obtained in (Curgus and Mascioni, 2004), and give explicit examples to further elaborate the ideas discussed. We then consider cubic polynomials, studying the results of (Frayer et al., 2014), and provide alternative proofs to some of their arguments. Keeping in spirit with the cubic case, we proceed to give a proof of Marden's theorem based on the award winning paper (Kalman, 2008). An interesting by-product of this discussion is that it paves way to the last chapter where an expository survey of Sendov's conjecture is discussed, proving the conjecture for $p(z)$ with roots on the unit circle. Furthermore, connections with the results discussed in the previous sections are then established.

Finally we remark on what has already been achieved, as well as a suggestion of an alternative route towards attacking the conjecture.

Formal definitions of all the results mentioned above are given in the following chapters, but generally, theorem statements and proofs have been customized to the context of this essay, hence may be stated differently from the original sources cited.

[1] This theorem is named after Morris Marden, however, it was proven 81 years earlier by Jorg Siebeck, hence in some literature it is often referred to as Siebeck's Lemma.

2. The Gauss-Lucas Theorem

Instead of having a section dedicated entirely to preliminaries, results, lemmas, definitions and examples will be stated and called as and when needed.

2.1 A proof of the Gauss-Lucas theorem

This section is based on the paper (Curgus and Mascioni, 2004). We begin with a proof of the Gauss-Lucas theorem, and give the preliminary definitions and results needed in the improvement of the theorem. We proceed to give a proof of the refined theorem, and conclude the chapter by looking at an example utilizing the new result.

Definition 2.1.1 (Convex set)**.** A set S is convex if and only if for all s_1, $s_2 \in S$, $\alpha s_1 + (1-\alpha)s_2 \in S$, for all $\alpha \in [0, 1]$

Definition 2.1.2 (Convex combination)**.** A convex combination of the points $s_1, s_2, \ldots, s_k$ is a point s of the form: $s = \sum_{j=1}^{k} \alpha_j s_j$ where $\sum_{j=1}^{k} \alpha_j = 1$ and each $\alpha_j \geq 0$.

Definition 2.1.3 (Convex hull)**.** The convex hull of a convex set S is the set of all convex combinations of points in S, that is:

$$H(S)^{1} := \Big\{ \sum_{j=1}^{k} \alpha_j s_j : s_j \in S,\ \alpha_j \geq 0,\ \sum_{j=1}^{k} \alpha_j = 1 \Big\}.$$

We now state and prove the theorem. There are several more aesthetic proofs of the Gauss-Lucas theorem, however, we prefer the following one since it introduces some useful notation early on which we will use in subsequent results.

Theorem 2.1.4 (Gauss-Lucas)**.** *The critical points of a non-constant complex polynomial $p(z)$ lie in the convex hull of the set of its roots.*

Proof. Let ζ be a non-trivial critical point of $p(z)$[2]. Let $z_1, z_2, \ldots, z_k$ be all the distinct roots of $p(z)$ with multiplicities $m_1, m_2, \ldots, m_k$, respectively. Since any given polynomial can be transformed to its monic equivalent by dividing each coefficient by the leading coefficient, without loss of generality we can write $p(z)$ in the form:

$$p(z) = \prod_{j=1}^{k} (z - z_j)^{m_j} \tag{2.1.1}$$

Taking logarithms of both sides yields:

$$\log(p(z)) = \sum_{j=1}^{k} m_j \log(z - z_j) \tag{2.1.2}$$

[1]Assuming the cardinality of S, $|S| = k$.
[2]this means $p'(\zeta) = 0$ but $p(\zeta) \neq 0$

Taking derivatives of both sides gives:

$$\frac{p'(z)}{p(z)} = \sum_{j=1}^{k} \frac{m_j}{z - z_j} \tag{2.1.3}$$

Recalling that $\bar{0} = 0$, we substitute ζ into 2.1.3 and obtain:

$$\begin{aligned} 0 = \frac{p'(\zeta)}{p(\zeta)} &= \sum_{j=1}^{k} \frac{m_j}{\zeta - z_j} = \sum_{j=1}^{k} \frac{m_j}{\bar{\zeta} - \bar{z}_j} \\ &= \sum_{j=1}^{k} \frac{m_j}{\bar{\zeta} - \bar{z}_j} \frac{|\zeta - z_j|^2}{|\zeta - z_j|^2} = \sum_{j=1}^{k} \frac{m_j}{|\zeta - z_j|^2}(\zeta - z_j) \end{aligned}$$

Thus we may write:

$$0 = \sum_{j=1}^{k} \frac{m_j}{|\zeta - z_j|^2}(\zeta - z_j) = \sum_{j=1}^{k} c_j(\zeta - z_j) \tag{2.1.4}$$

where:

$$c_j := \frac{m_j}{|\zeta - z_j|^2} > 0 \quad j = 1, \dots, k \tag{2.1.5}$$

from 2.1.4 we solve for ζ to get:

$$\zeta = \frac{1}{\sum_{j=1}^{k} c_j} \left(\sum_{j=1}^{k} c_j z_j \right) = \sum_{j=1}^{k} d_j z_j \tag{2.1.6}$$

where:

$$d_j := \frac{c_j}{\sum_{\mu=1}^{k} c_\mu}, \quad j = 1, \dots, k. \tag{2.1.7}$$

We note: $0 < d_j \leq 1$, for $j = 1, \dots, k$, and $\sum_{j=1}^{k} d_j = 1$. Hence 2.1.6 shows that ζ is in the convex hull of the roots of $p(z)$ hence the proof is done.

□

2.2 Refining the Gauss-Lucas theorem

By the Gauss-Lucas theorem, a critical point c of $p(z)$ is on the boundary of the convex hull of the zeros of $p(z)$ if and only if c is a repeated root. From this, we can intuitively deduce that the nontrivial critical points are confined to a smaller region, strictly inside the convex hull of the roots. Our main result in this section is simply to characterise or describe this new region explicitly.

Definition 2.2.1. Let $Z(p)$=$\{z \in \mathbb{C} : p(z) = 0\}$, that is, the set of all roots of $p(z)$.

Definition 2.2.2 (Lucas Polygon)**.** We define the Lucas polygon as the boundary of the convex hull of $Z(p)$.

Definition 2.2.3. The closed disc in $\mathbb{C}$ with radius $r > 0$ and center $u \in \mathbb{C}$ is denoted by $D(u, r)$.

Notations of some quantities needed in the refinement proof are given below:

Definition 2.2.4. [3] Let $Z(p)$ be as defined above, then:

$$\begin{aligned}
\omega(p, w) &:= \min\{|w - v| : v \in Z(p), w \neq v\}, w \in Z(p), \\
\omega(p) &:= \min\{|w - v| : w, v \in Z(p), w \neq v\}, \\
\Omega(p) &:= \max\{|w - v| : w, v \in Z(p), w \neq v\}, \\
\tau(p, w) &:= \min\{|w - v| : v \in Z(p'), w \neq v\}, w \in Z(p), \\
\tau(p) &:= \min\{|w - v| : w \in Z(p), v \in Z(p'), w \neq v\}, \\
T(p) &:= \max\{|w - v| : w \in Z(p), v \in Z(p'), w \neq v\}.
\end{aligned}$$

We refer the reader to (Curgus and Mascioni, 2003) for the proofs of the following propositions. Let m_w be the multiplicity of $w \in Z(p)$, then:

Proposition 2.2.5. $0 < \frac{1}{n}\omega(p) \leq \tau(p) \leq \frac{1}{2\sin(\pi/n)}\omega(p)$

Proposition 2.2.6. $0 < \frac{m_w}{n}\omega(p, w) \leq \tau(p, w) \leq \frac{1}{2\sin((\pi/n)-m_w)}\omega(p, w)$

We also need the following result:

Proposition 2.2.7. $T(p) \leq \sqrt{\Omega(p)^2 - \tau(p)^2}$

Proof. Let $w \in Z(p)$ and $\zeta \in Z(p')$, $w \neq \zeta$, such that $|w - \zeta| = T(p)$. Suppose a line l passes through ζ and perpendicular to the segment $\overline{w\zeta}$, then l separates the $\mathbb{C} - plane$ into two half planes. Since all critical points of $p(z)$ lie in the convex hull of $Z(p)$ by Gauss-Lucas theorem, we can find some $v \in Z(p)$ in the half plane not containing w. We can thus deduce that the angle between the line segments $\overline{\zeta w}$ and $\overline{\zeta v}$ is greater than $\frac{\pi}{2}$. Equivalently:

$$\Omega^2 \geq |w - v|^2 \geq T(p)^2 + \tau(p)^2 \tag{2.2.1}$$

The result then follows from equation (2.2.1). □

[3] We consider $p(z)$ to be of degree $n \geq 2$, and assume there are at least two distinct roots.

Remark 2.2.8. Using propositions 2.2.5 and 2.2.7 we get:

$$T(p) \leq \sqrt{\Omega(p)^2 - \omega(p)^2/n^2} \tag{2.2.2}$$

We are now ready to state and prove the main result for this section.

Theorem 2.2.9 (A contraction of the Lucas polygon)**.** *(Curgus and Mascioni, 2004) Let $p(z) := \prod_{j=1}^{k}(z - w_j)^{m_j}$ be a non-constant polynomial of degree $n = \sum_{j=1}^{k} m_j$. Define $\delta = \sum_{j=1}^{k} \delta_j$, where*

$$\delta_j := \frac{m_j}{m_j + (n^2\Omega(p)^2 - \omega(p)^2)\sum_{i\neq j}\frac{1}{m_i\omega(p,w_i)^2}}$$

and define

$$w_b := \frac{1}{\delta}\sum_{j=1}^{k}\delta_j w_j.$$

Then $\delta \in (0,1)$ and all the non-trivial critical points of $p(z)$ lie in the new region defined by contracting the Lucas polygon of $p(z)$, with contraction coefficient $1-\delta$, centered at w_b, and removing from the region the interiors of the discs $D(w_j, m_j\omega(p,w_j)/n)$, $j = 1,\ldots,k,$.

Proof. We note that, the contraction centered at w_b with contraction coefficient $1-\delta$ is an affine bijective transformation of $\mathbb{C}$, hence the Lucas polygon of p is mapped onto the convex hull of the points

$$v_j := \delta w_b + (1-\delta)w_j, j = 1,\ldots,k.$$

Our task is to show that all the roots of p are contained in this smaller polygon. We achieve this by giving detailed estimates for the coefficients c_j and d_j, $j = 1,\ldots,k$ as defined in the proof of 2.1.4, the Gauss-Lucas theorem. For ease of notation, let $\omega(p) = \omega$ and $\Omega(p) = \Omega$. Hence, for $j = 1,\ldots,k$:

$$d_j = \frac{c_j}{\sum_{\mu=1}^{k} c_\mu} = \frac{m_j}{|\zeta - w_j|^2}\frac{1}{\sum_{\mu=1}^{k} c_\mu} \tag{2.2.3}$$

$$= \frac{m_j}{n}\left[\frac{n}{|\zeta - w_j|^2\sum_{\mu=1}^{k} c_\mu}\right] \tag{2.2.4}$$

$$= \frac{m_j}{n}\left[\frac{n}{m_j + |\zeta - w_j|^2\sum_{i\neq j}\frac{m_i}{|\zeta - w_i|^2}}\right] \tag{2.2.5}$$

By remark 2.2.8 and proposition 2.2.6 we get

$$|\zeta - w_j| \leq T(p) \leq \sqrt{\Omega(p)^2 - \omega(p)^2/n^2}$$

and

$$|\zeta - w_j| \geq \tau(p, w_i) \geq \frac{m_i \omega(p, w_i)}{n}$$

respectively. Hence equality 2.2.5 implies:

$$d_j \geq \frac{m_j}{m_j + (n^2 \Omega(p)^2 - \omega(p)^2) \sum_{i \neq j} \frac{1}{m_i \omega(p,w_i)^2}} =: \delta_j. \tag{2.2.6}$$

Thus, by the above result 2.2.6 and our definition of d_j from 2.2.5 we get that:

$$0 < \delta_j \leq d_j \leq 1, \quad j = 1, \ldots, k \tag{2.2.7}$$

and $\sum_{\mu=1}^{k} \delta_\mu = \delta < 1$. Furthermore, for arbitrary j we have:

$$^4 d_j = 1 - \sum_{\substack{\mu=1 \\ \mu \neq j}}^{k} d_\mu \leq 1 - \sum_{\substack{\mu=1 \\ \mu \neq j}}^{k} \delta_\mu = 1 - \delta + \delta_j < 1. \tag{2.2.8}$$

Keeping in mind the definition of ζ from 2.1.6, combining with 2.2.6 and 2.2.7 above, we conclude that all the non-trivial roots of p' lie in the convex region defined by:

$$R_w := \left\{ \sum_{\mu=1}^{k} t_\mu w_\mu : \delta_\mu \leq t_\mu \leq 1 - \delta + \delta_\mu, \sum_{\mu=1}^{k} t_\mu = 1 \right\}.$$

Our goal now is to show that R_w is indeed the convex hull of the points bounded by the contracted polygon. Hence, let R_v be the convex hull of the points:

$$v_j := \delta w_b + (1-\delta) w_j = \sum_{\substack{\mu=1 \\ \mu \neq j}}^{k} \delta_\mu w_\mu + (1 - \delta + \delta_j) w_j, j = 1, \ldots, k$$

Or more succinctly,

$$R_v := \left\{ \sum_{\mu=1}^{k} s_\mu v_\mu : 0 \leq s_\mu \leq 1, \sum_{\mu=1}^{k} s_\mu = 1 \right\}.$$

We then need to establish that $R_v = R_w$. We proceed by constructing two simplices S_v and S_w respectively where:

$$S_v = \left\{ \begin{bmatrix} s_1 & s_2 & s_3 & \ldots & s_{k-1} & s_k \end{bmatrix}^T : 0 \leq s_\mu \leq 1, \sum_{\mu=1}^{k} s_\mu = 1 \right\}$$

[4] Recall that $\sum_{j=1}^{k} d_j = 1$. Combining with 2.2.7 the result follows.

and

$$S_w = \left\{ [t_1 \quad t_2 \quad t_3 \quad \dots \quad t_{k-1} \quad t_k]^T : \delta_\mu \le t_\mu \le 1 - \delta + \delta_\mu, \sum_{\mu=1}^{k} t_\mu = 1 \right\}$$

A one-to-one correspondence between the simplices is established such that:

$$\sum_{\mu=1}^{k} t_\mu w_\mu = \sum_{\mu=1}^{k} s_\mu v_\mu$$

Letting $\mathbf{t} = [t_1 \quad t_2 \quad t_3 \quad \dots \quad t_{k-1} \quad t_k]^T$, and $\mathbf{s} = [s_1 \quad s_2 \quad s_3 \quad \dots \quad s_{k-1} \quad s_k]^T$, we achieve this by solving the system $\mathbf{As} = \mathbf{t}$, where the matrix $\mathbf{A}$ is given by:

$$\mathbf{A} = \begin{pmatrix} 1-\delta+\delta_1 & \delta_1 & \delta_1 & \cdots & \delta_1 & \delta_1 \\ \delta_2 & 1-\delta+\delta_2 & \delta_2 & \cdots & \delta_2 & \delta_2 \\ \delta_3 & \delta_3 & 1-\delta+\delta_3 & \cdots & \delta_3 & \delta_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \delta_{k-1} & \delta_{k-1} & \delta_{k-1} & \cdots & 1-\delta+\delta_{k-1} & \delta_{k-1} \\ \delta_k & \delta_k & \delta_k & \cdots & \delta_k & 1-\delta+\delta_k \end{pmatrix}$$

Solving the system, with $\mathbf{A}^{-1}$ given by:

$$\mathbf{A}^{-1} = \frac{1}{1-\delta} \begin{pmatrix} 1-\delta_1 & -\delta_1 & -\delta_1 & \cdots & -\delta_1 & -\delta_1 \\ -\delta_2 & 1-\delta_2 & -\delta_2 & \cdots & -\delta_2 & -\delta_2 \\ -\delta_3 & -\delta_3 & 1-\delta_3 & \cdots & -\delta_3 & -\delta_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\delta_{k-1} & -\delta_{k-1} & -\delta_{k-1} & \cdots & 1-\delta_{k-1} & -\delta_{k-1} \\ -\delta_k & -\delta_k & -\delta_k & \cdots & -\delta_k & 1-\delta_k \end{pmatrix}$$

Thus for the simplex $\mathbf{t} \in S_w$, the entries of $\mathbf{s} = \mathbf{A}^{-1}\mathbf{t}$ are given by:

$$s_j = \frac{t_j - \delta_j}{1-\delta}, \quad j = 1, \dots, k \tag{2.2.9}$$

Equation 2.2.9 above implies that we have the injective mapping $\mathbf{A}^{-1} : S_w \hookrightarrow S_v$. Conversely, the entries of $\mathbf{t} = \mathbf{As}$ are given by:

$$t_j = (1-\delta)s_j + \delta_j, \quad j = 1, \dots, k$$

This gives us another injective mapping $\mathbf{A} : S_v \hookrightarrow S_w$. Our one-to-one correspondence is thus established.

We note that, this verifies that all non-trivial roots of $p'(z)$ lie in the convex hull R_v. Furthermore, this region is the contraction of the Lucas polygon of $p(z)$ with contraction coefficient $1-\delta$ and centered at w_b. We now need to justify the part of our statement about removing discs centered at w_j.

From definition 2.2.4, we defined $\tau(p, w)$ to be the minimum distance from the root w to the nearest critical point v of $p(z)$. This implies that none of the discs $D(w, \tau(p, w))$ contains a non-trivial root of $p'(z)$. By inequality 2.2.6, this is still true for the smaller regions $D(w, m_w\omega(p, w)/n)$. This completes the proof. □

This seemingly abstract result will be further clarified through the example we will construct below:

We consider $p(z) = \prod_{j=1}^{n}(z - r_j)$, with $|r_j| = 1$ for each $j = 1, \ldots, n$, that is, polynomials whose roots are on the unit circle.

Example 2.2.10. Let $\varphi = \frac{1}{2} + i\frac{\sqrt{3}}{2}$, then $\overline{\varphi} = \frac{1}{2} - i\frac{\sqrt{3}}{2}$. Consider the polynomial:

$$\begin{aligned} p(z) &= (z - 1)(z - \varphi)(z - \overline{\varphi}) \\ &= z^3 - 2z^2 + 2z - 1 \end{aligned}$$

then:

$$p'(z) = 3z^2 - 4z + 2$$

Hence the critical points of $p(z)$ are:

$$c_1 = \frac{2}{3} - \frac{\sqrt{2}}{3}i \quad \text{and} \quad c_2 = \frac{2}{3} + \frac{\sqrt{2}}{3}i$$

The multiplicity of each root, $m_j = 1$, $j = 1, 2, 3$. We need to compute δ and w_b.

By definition 2.2.4:

$$\Omega(p) = |\varphi - \overline{\varphi}| = |\sqrt{3}i| = \sqrt{3}.$$

Similarly:

$$\omega(p) = |1 - \varphi| = \omega(p, 1) = \omega(p, \varphi) = \omega(p, \overline{\varphi}) = 1,$$

Hence each δ_j satisfies:

$$\delta_j = \frac{1}{1 + (n^2\Omega^2 - \omega^2)(1 + 1)} = \frac{1}{1 + 2(9 \times 3 - 1)} = \frac{1}{53}$$

Therefore $\delta = \delta_1 + \delta_2 + \delta_3 = \frac{3}{53}$, hence $1 - \delta = \frac{50}{53}$.

Recall that $w_b := \frac{1}{\delta}\sum_{j=1}^{k} \delta_j w_j$, hence:

$$w_b = \frac{53}{3}\left(\frac{1}{53} + \frac{\varphi}{53} + \frac{\overline{\varphi}}{53}\right) = \frac{1}{3}(1 + \varphi + \overline{\varphi}) = \frac{2}{3}.$$

Hence the nontrivial critical points of $p(z)$ lie inside the new region obtained by contracting the triangle $\triangle 1\varphi\overline{\varphi}$ by a factor of $\frac{50}{53}$, and shaving off the interiors of the discs: $D(1,\frac{1}{3}), D(\varphi,\frac{1}{3})$ and $D(\overline{\varphi},\frac{1}{3})$. This is the shaded region depicted in the figure below.

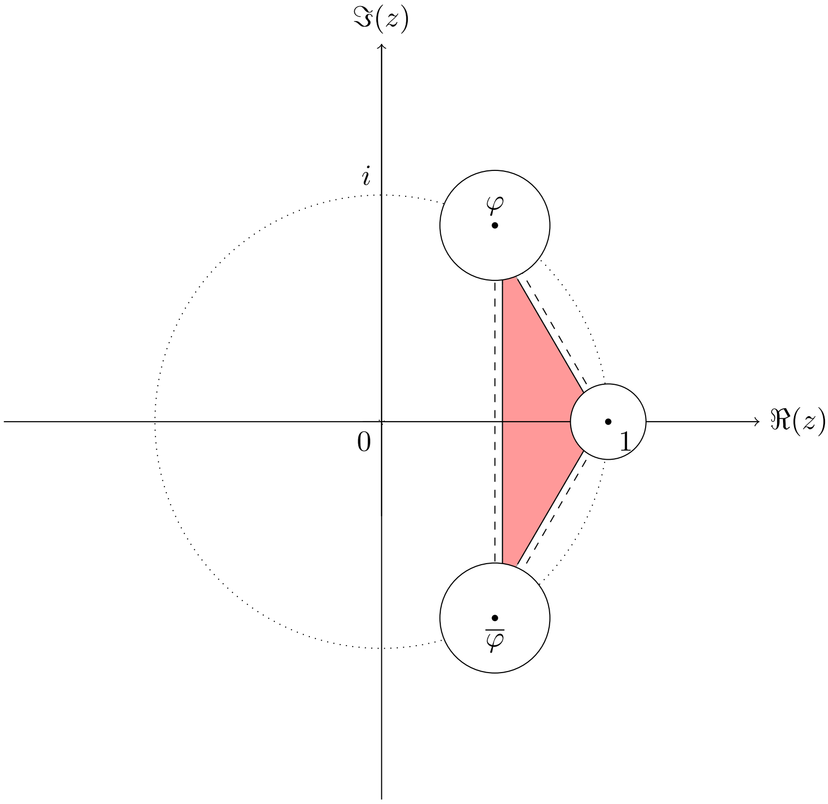

Figure 2.1: Illustrating the region containing the critical points.

3. Cubic Polynomials and Marden's Theorem

This section is mainly concerned with the study of cubic polynomials. This particular class of polynomials is easier to study, hence it is by no means surprising that many results that later became general theorems about polynomials started off as intuitive ideas about geometrical relations between the roots and the critical points of cubic polynomials.

Conversely, unexpected and interesting geometrical properties are often unravelled when we specialize the often abstract general results about polynomials to the particular case of cubic polynomials. Our discussion is based on the papers by (Frayer et al., 2014) and (Kalman, 2008).

Our purpose here is twofold: we give a discussion of results that pave way to the objective of the next chapter, and we also discuss Marden's theorem, a more accurate refinement of the Gauss-Lucas theorem for cubic polynomials.

3.1 The Geometry of Cubic Polynomials

Remark 3.1.1. It is important to note that given a triangle formed from the roots of $p(z)$, we may, without loss of generality, rotate, scale, and translate it in a manner convenient to our goal. More formally, we can define a linear function $T : \mathbb{C} \longrightarrow \mathbb{C}$ by $T(z) = \alpha z + \beta$, with $\alpha = re^{i\theta} \neq 0$, through which we impose the transformations. We demonstrate this assertion explicitly when we get to the section on Marden's theorem.

Bearing the above remark in mind, we restrict the discussion in this section to cubic polynomials whose roots are on the unit circle, defined below.

Definition 3.1.2. By Γ we denote the family of cubic polynomials $p : \mathbb{C} \longrightarrow \mathbb{C}$ such that:

$$p(z) = (z-1)(z-r_1)(z-r_2) \text{ with } |r_1| = 1 = |r_2|.$$

Definition 3.1.3 (centroid of a triangle)**.** This is the point of intersection of the medians of the triangle. A useful property of the centroid is that it divides each median in the ratio $2 : 1$, the longest side being the one closest to the vertex.

Definition 3.1.4 (center of $p(z)$)**.** Let $p \in \Gamma$ and $c \in \mathbb{C}$, then c is the center of $p(z)$ if $p''(c) = 0$. Geometrically, this is the centroid of the triangle whose vertices are $1, r_1$ and r_2.

Remark 3.1.5. We note that, every $p \in \Gamma$ has a unique center, the following result gives a converse of this fact.

Theorem 3.1.6 (Characterising a polynomial by the center:)**.** *Let $c \in \mathbb{C}$, then:*

- *$p \in \Gamma$ has center $\frac{1}{3}$ if and only if $p(z) = (z-1)(z^2 - r^2)$ for some r with $|r| = 1$.*
- *If $0 < |c - \frac{1}{3}| \leq \frac{2}{3}$, then c determines $p \in \Gamma$ uniquely.*
- *If $|c - \frac{1}{3}| > \frac{2}{3}$, then there is no $p \in \Gamma$ with center c.*

Proof. We give a geometrical argument:

Let c be the center of $p(z) \in \Gamma$. Then c is the centroid of $\triangle 1 r_1 r_2$, where r_1 and r_2 are unknown. Denote the midpoint of side $\overline{r_1 r_2}$ by w, which is still in the closed unit disc.

We note that, the segment $\overline{1w}$ is a median of $\triangle 1r_1r_2$, hence by the property relating the centroid and median, $c = \frac{2}{3}(w) + \frac{1}{3}(1)$. It follows immediately that $|c - \frac{1}{3}| \leq \frac{2}{3}$.

Solving for w gives $w = \frac{3c-1}{2}$. Suppose $c \neq \frac{1}{3}$, (hence $w \neq 0$). $\overline{r_1r_2}$ is a chord of the unit circle [1], hence the perpendicular bisector passes through the center of the circle 0 and the midpoint w.

Construct the line l through w and perpendicular to $\overline{0w}$. Since w is in the closed unit disc, l will intersect the unit circle in two points: these are r_1 and r_2.

However, if $c = \frac{1}{3}$, then $w = 0$. This means the midpoint of $\overline{r_1r_2}$ is 0, hence $r_1 = -r_2$. Conversely, if $r_1 = -r_2$, then $p(z) = (z-1)(z-r_1)(z+r_1) = z^3 - z^2 - r_1^2 z + r_1^2$. Then $p''(z) = 6z - 2$, and the center $c = \frac{1}{3}$. □

The figure below is an illustration of the construction of r_1 and r_2.

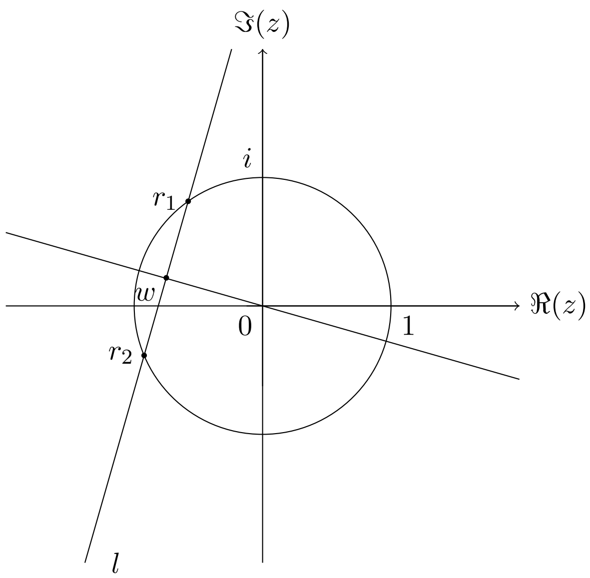

Figure 3.1: Illustrating the construction of r_1 and r_2

Remark 3.1.7. Armed with the result above, a natural question to ask is: can we find an analogous characterisation of $p \in \Gamma$ by critical points? It turns out that indeed we can.

Definition 3.1.8 (Inverse of a point P with respect to an inversion circle I)**.** Let I be a circle with center O and radius $r > 0$. Let P be the point to be inverted with respect to I, then P', the image of P, satisfies: $|OP||OP'| = r^2$, with P' lying on the ray through $\overline{OP}$.

For the proof of the following proposition, we refer to Chapter 3 of (Needham, 1999).

Proposition 3.1.9. The inverse of a circle of radius a and center (x, y) with respect to an inversion circle I with inversion center (x_0, y_0) and inversion radius r is another circle with center:

[1] since $|r_1| = 1 = |r_2|$

$$x' = x_0 + s(x - x_0)$$
$$y' = y_o + s(y - y_0)$$

and radius $r' = |s|a$, where:

$$s = \frac{r^2}{(x - x_0)^2 + (y - y_0)^2 - a^2}.$$

Definition 3.1.10. Let $\alpha > 0$. Denote by C_α the circle of diameter α that passes through 1 and $1 - \alpha$ in the complex plane. That is:

$$C_\alpha = \left\{ z \in \mathbb{C} : \left| z - \left(1 - \frac{\alpha}{2}\right) \right| = \frac{\alpha}{2} \right\}.$$

Theorem 3.1.11. *Let $z \in \mathbb{C}$ with $Re(z) < 1$. Then $z \in C_\alpha$ if and only if*

$$Re\left(\tfrac{1}{1-z}\right) = \tfrac{1}{\alpha}.$$

For a short and concise proof of the theorem, see (Frayer et al., 2014). However, we prove a more general result which allows us to explicitly calculate α.

Proposition 3.1.12. Let $a \neq 0$ be real and $z \in \mathbb{C}$ with $z \neq a$, then:

$$Re\left(\frac{1}{a - z}\right) = \frac{1}{2a} - \frac{|z|^2 - a^2}{2a|a - z|^2}.$$

Proof. Let $z = x - iy$, then:

$$\frac{1}{a - z} = \frac{1}{(a - x) + iy} = \frac{(a - x) - iy}{(a - x)^2 + y^2}$$

Hence

$$Re\left(\frac{1}{a - z}\right) = \frac{(a - x)}{(a - x)^2 + y^2}$$

On the other hand:

$$\begin{aligned} \frac{1}{2a} - \frac{|z|^2 - a^2}{2a|a - z|^2} &= \frac{|a - z|^2 + a^2 - |z|^2}{2a|a - z|^2} \\ &= \frac{(a - x)^2 + y^2 + a^2 - (x^2 + y^2)}{2a((a - x)^2 + y^2)} \\ &= \frac{(a - x)}{(a - x)^2 + y^2} \end{aligned}$$

This completes the proof. □

With the above result we derive several corollaries:

Corollary 3.1.13 (explicit computation of α)**.** In accordance with the statement of theorem 3.1.11, suppose $a = 1$ and $Re(z) < 1$, then:

$$Re\left(\frac{1}{1-z}\right) = \frac{(1-x)}{(1-x)^2 + y^2} = \frac{1}{\left(\frac{(1-x)^2+y^2}{(1-x)}\right)}$$

Hence, for $1 \neq z = x + iy \in \mathbb{C}$, $z \in C_\alpha$ where:

$$\alpha = \frac{(1-x)^2 + y^2}{(1-x)}$$

Corollary 3.1.14. If $a = 1$, and $1 \neq z \in \mathbb{C}$ such that $|z| = 1$, then $Re\left(\frac{1}{1-z}\right) = \frac{1}{2}$. Or equivalently, $\alpha = 2$.

And more generally:

Corollary 3.1.15. The complex number $z \neq 1$ lies in the closed unit disc if and only if there is a unique $\alpha \in (0, 2]$ for which z lies on C_α.

Example 3.1.16. For the case where a critical point $c \in \mathbb{C}$ of $p(z) = (z-1)(z-r_1)(z-r_2)$ lies on the unit circle, by Gauss-Lucas theorem, c will coincide with one of the roots of $p(z)$, that is:

$$c \in \{1, r_1, r_2\}.$$

If $c = 1$, then $p(z) = (z-1)^2(z-r)$ for some $r \in \mathbb{C}$ with $|r| = 1$.

Otherwise $p(z) = (z-1)(z-r)^2$. This means $p'(z) = 3(z-r)\left(z - \frac{r+2}{3}\right)$. Thus the second critical point will be $c_2 = \frac{r+2}{3} \in T_{\frac{2}{3}}$.

Theorem 3.1.17. *Let $p(z) = (z-1)(z-z_1)\cdots(z-z_n)$, where $z_k = e^{i\theta_k}$ for each k. Let $c_1, \ldots, c_n$ denote the critical points of $p(z)$, and suppose that $1 \neq c_k \in T_{\alpha_k}$ for each k. Then:*

$$\sum_{k=1}^{n} \frac{1}{1-c_k} = 2\sum_{k=1}^{n} \frac{1}{1-z_k} \tag{3.1.1}$$

and

$$\sum_{k=1}^{n} \frac{1}{\alpha_k} = n. \tag{3.1.2}$$

Proof. We defer the proof of equality 3.1.1 until the next chapter where we prove a more general result in the discussion of Sendov's conjecture.

We note, however, that with equality 3.1.1 established, by corollary 3.1.14, we obtain:

$$\sum_{k=1}^{n} \frac{1}{\alpha_k} = Re\left(\sum_{k=1}^{n} \frac{1}{1-c_k}\right) = 2\sum_{k=1}^{n} Re\left(\frac{1}{1-z_k}\right) = 2\sum_{k=1}^{n} \frac{1}{2} = n. \tag{3.1.3}$$

□

Corollary 3.1.18. Let $p \in \Gamma$, and let $c_1 \neq 1$, $c_2 \neq 1$ be the critical points of p. If $c_1 \in C_\alpha$ and $c_2 \in C_\beta$, then:

$$\frac{1}{\alpha} + \frac{1}{\beta} = 2.$$

Remark 3.1.19. We claim that geometrically, corollary 3.1.18 can be interpreted as saying C_α is the inversion of C_β with respect to C_1. We justify this claim below:

Proof. We note that C_1 is centered at $(x_0, y_0) = (\frac{1}{2}, 0)$ with radius $r = \frac{1}{2}$.

Suppose corollary 3.1.18 holds, and let $c_1 \in C_\alpha$ and $c_2 \in C_\beta$. This means $\beta > 0$, hence $\alpha > \frac{1}{2}$, and by definition 3.1.10, C_α has center $(x, y) = (1 - \frac{\alpha}{2})$ and radius $a = \frac{\alpha}{2}$.

We determine the inversion of C_α with respect to C_1. By proposition 3.1.9:

$$s = \frac{r^2}{(x-x_0)^2 + (y-y_0)^2 - a^2} = \frac{1}{1-2\alpha}$$

Hence, the inversion has radius:

$$r' = |s|a = \left|\frac{1}{1-2\alpha}\right| \frac{\alpha}{2} = \frac{\alpha}{2(2\alpha-1)}$$

and center:

$$\begin{aligned}(x_0 + s(x-x_0), y_0 + s(y-y_0)) &= \left(\frac{1}{2} + \frac{1}{1-2\alpha}\left(\frac{1}{2} - \frac{\alpha}{2}\right), 0\right) \\ &= \left(1 + \frac{1}{2}\left(\frac{1-\alpha}{1-2\alpha} - 1\right), 0\right) \\ &= (1 - r', 0).\end{aligned}$$

This is the inversion C_β with $\beta = 2r' = \frac{\alpha}{2\alpha-1}$. Hence:

$$\frac{1}{\alpha} + \frac{1}{\beta} = \frac{1}{\alpha} + \frac{2\alpha-1}{\alpha} = 2$$

as required.

□

Theorem 3.1.20. *(Frayer et al., 2014) No polynomial $p \in \Gamma$ has a critical point strictly inside $C_{2/3}$.*

Proof. Suppose, for the sake of contradiction that $c \in \mathbb{C}$ is a critical point of some polynomial $p \in \Gamma$, and that c lies strictly in $C_{2/3}$.

This implies that c lies on C_α for some $\alpha \in (0, \frac{2}{3})$. Then the other critical point of p lies on C_β, by corollary 3.1.18:

$$\frac{1}{\beta} = 2 - \frac{1}{\alpha} < 2 - \frac{3}{2} = \frac{1}{2}.$$

However, this implies that $\beta > 2$, which contradicts corollary 3.1.15. This completes the proof. □

We are now in a position to state the result characterising polynomials by their critical points.

Theorem 3.1.21. *(Frayer et al., 2014). Let $c \in \mathbb{C}$*

- *if $c \notin \{1, \frac{1}{3}\}$, there is at most one $p \in \Gamma$ with a critical point at c.*
- *if c lies strictly inside $C_{2/3}$, or strictly outside C_2, then there is no $p \in \Gamma$ with a critical point at c.*
- *$p \in \Gamma$ has a critical point at 1 if and only if $p(z) = (z-1)^2(z-r)$ for some r on the unit circle.*
- *$p \in \Gamma$ has a critical point at $-\frac{1}{3}$ if and only if $p(z) = (z-1)(z-r)\left(z + \frac{5r+3}{3r+5}\right)$ for some r on the unit circle.*

Remark 3.1.22. For a full treatment of theorem 3.1.21, see (Frayer et al., 2014). However, the reader should note that the second bullet point follows from theorem 3.1.20 and the Gauss-Lucas theorem [2].

The third bullet point was proven as example 3.1.16.

We end this section with a figure illustrating the regions described in theorems 3.1.6 and 3.1.21: no critical point of $p \in \Gamma$ can be found in the purple region, and no center lies outside the red region.

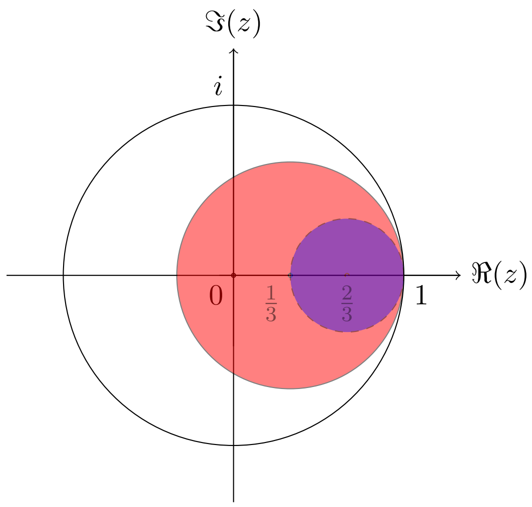

Figure 3.2: Illustrating theorems 3.1.21 and 3.1.6

[2]By the Gauss-Lucas theorem, all critical points lie in the convex hull of the roots of p hence in the unit disc, thus if c lies outside C_2, then $p \notin \Gamma$.

3.2 Marden's Theorem

In this section we study Marden's theorem, giving its proof and concluding with a short survey of the possible generalisations. As purported in the previous section, we begin this section with a proof justifying remark 3.1.1, stated more formally below:

Proposition 3.2.1. (Minda and Phelps, 2008) Let $p(z) = (z - z_1)(z - z_2)(z - z_3)$ with critical points r_1, r_2. Define a linear transformation $T(z) = \alpha z + \beta$, where $\alpha \neq 0$, and β are arbitrary complex numbers. Then the critical points of the polynomial $P_T(z) = (z - T(z_1))(z - T(z_2))(z - T(z_3))$ are equal to $T(r_1)$ and $T(r_2)$.

Proof. We substitute $T(z)$ into $P_T(z)$ to get:

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

Taking derivatives both sides yields:

$$\alpha P_T'(T(z)) = \alpha^3 p'(z)$$

Hence $P_T'(T(r_1)) = P_T'(T(r_2)) = 0$.

□

Remark 3.2.2. The above proposition shows that T carries the roots of p' to the roots of P_T'. Thus, a suitably chosen linear map T allows us to scale, rotate, and translate the roots of p to more convenient positions in the complex plane, hence, a proof of the theorem using a simpler polynomial P_T suffices for the general case.

The following theorem of Steiner gives us a definition of the so-called Steiner inellipse of a triangle.

Theorem 3.2.3 (Steiner)**.** *Every triangle admits a unique inscribed ellipse which is tangent to the sides of the triangle at their midpoints. If z_1, z_2, z_3 are the vertices of the triangle, then the foci of this ellipse are:*

$$g \pm \sqrt{g^2 - \frac{1}{3}(z_1 z_2 + z_2 z_3 + z_1 z_3)}$$

where $g = \frac{1}{3}(z_1 + z_2 + z_3)$.

Definition 3.2.4 (Steiner Inellipse)**.** The ellipse inscribed in $\triangle z_1 z_2 z_3$ and tangent to the midpoints of the sides is called the Steiner inellipse.

The figure below illustrates the geometric configuration of the Steiner inellipse for a given triangle:

Here z_1, z_2, z_3 denote the roots of $p(z)$, F_1, F_2 are the foci of the ellipse and α, β, γ are the points of tangency, which are also the midpoints of the sides of $\triangle z_1 z_2 z_3$.

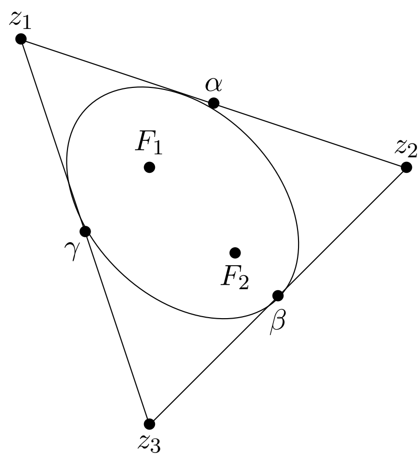

Figure 3.3: Steiner inellipse

Remark 3.2.5. A corollary of theorem 3.2.3 was given by Siebeck, now commonly referred to as Marden's theorem, whose formal statement and proof we give below.

Theorem 3.2.6 (Marden)**.** *Let $p(z)$ be a cubic complex polynomial with noncollinear zeros z_1, z_2, z_3, then the critical points of $p(z)$ are the foci of the Steiner inellipse of the triangle whose vertices are z_1, z_2, z_3.*

We note that, when stated in this form, the proof of Marden's theorem is essentially the verification of theorem 3.2.3, that is, establishing the existence and uniqueness of the Steiner inellipse. This is the approach that we adopt for our proof.

The following lemma (one of the optical properties of an ellipse) will be useful for our proof, hereby stated without proof, but the reader may consult (Kalman, 2008) for a geometric proof.

Lemma 3.2.7 (An ellipse, external point and tangent lines)**.** Let F_1 and F_2 be the foci of an ellipse $\mathcal{E}$, and A a point external to the ellipse, hence there are two lines through A that are tangent to $\mathcal{E}$. Let G_1 and G_2 be the points of tangency of these to lines, then $\angle F_1AG_1 = \angle F_2AG_2$.

The figure below gives the geometrical configuration of the statement of the above lemma.

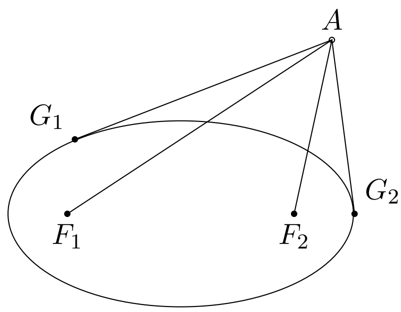

Figure 3.4: Ellipse, external point and tangent lines.

A variant of the following lemma (stated here differently) was given by Marden.

Before stating the lemma let us recall the following quick fact from high school algebra:

Proposition 3.2.8. If α and β are the zeros of $y = ax^2 + bx + c$, then $\alpha + \beta = -\frac{b}{a}$ and $\alpha\beta = \frac{c}{a}$.

Lemma 3.2.9. Let $T = \triangle z_1z_2z_3$ be the triangle formed from the roots of $p(z)$. If $\mathcal{E}$ is the ellipse whose foci are the critical points of $p(z)$, and $\mathcal{E}$ passes through the midpoint of one side of T, then the ellipse is tangent to that side of T.

Proof. By proposition 3.2.1, we choose the vertices of our triangle (and hence the roots of $P_T(z)$ to be 1, -1 and $w = a + ib$, $b > 0$. We analyse the ellipse passing through zero, the midpoint of the side lying on the real axis.

To show that the ellipse is tangent to this side, it suffices to show that the lines from each foci to the origin subtend equal angles with the real axis. In fact, given this configuration, the result is a consequence of the optical property of ellipses.

Since we know the roots of $p(z)$, we write it as:

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

Taking derivatives, we get:

$$p'(z) = 3z^2 - 2zw - 1 = 3\left(z^2 - \frac{2w}{3} - \frac{1}{3}\right).$$

Let $r_1 = re^{i\theta_1}$ and $r_2 = se^{i\theta_2}$ be the critical points of $p(z)$ (where $0 \leq \theta_1,\ \theta_2 < 2\pi$). Then by proposition 3.2.8 we get:

$$r_1 + r_2 = \frac{2}{3}w \quad \text{and} \quad r_1r_2 = -\frac{1}{3}$$

From the first of the above equalities, we deduce that at least one of the critical points of p lies in the upper half-plane. The second equality shows that $\theta_1 + \theta_2 = \pi$, thus both critical points lie in the upper half-plane. Knowingly blurring the distinction between vectors and complex numbers, we note that, the angles subtended by the vectors r_1 and r_2 with the positive $x - axis$ add up to π radians.

Hence, either both critical points are on the imaginary axis, or one makes an acute angle with the positive real axis whilst the other makes an equal angle with the negative real axis.

Hence, the lines from the foci of $\mathcal{E}$ to 0 make equal angles with the real axis. Thus $\mathcal{E}$ is tangent to the real axis as required. □

The following lemma, which is an extension of lemma 3.2.9 above, completes the existence part of our approach.

Lemma 3.2.10. Given the same configuration of the ellipse $\mathcal{E}$ as in lemma 3.2.9 above, that is, $\mathcal{E}$ being tangent to one side of triangle T, then $\mathcal{E}$ is tangent to all three sides of T.

Proof. (Sketch) The reasoning behind this proof is very much similar to the idea behind the proof of the previous lemma, hence we just sketch it here.

Appealing to proposition 3.2.1 once more, we choose the vertices of our triangle to be 0, 1 and $w = a+bi$ where $b > 0$. The ellipse $\mathcal{E}$ is chosen to be tangent to the side along the real axis, hence it is tangent at $x = \frac{1}{2}$.

Our goal is to show that it is also tangent to the side $\overline{0w}$. With our choice of vertices, the corresponding polynomial $p(z)$ is:

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

Taking derivatives, we get:

$$p'(z) = 3z^2 - 2(1+w)z + w = 3\left(z^2 - \frac{2w}{3} - \frac{1}{3}\right).$$

As in the previous lemma, our critical points r_1 and r_2 satisfy:

$$r_1 + r_2 = \frac{2}{3}(1+w)$$

From the above equation, as well as knowing that the critical points are the foci of an ellipse which is tangent to the real axis, we deduce that both critical points r_1 and r_2 are in the upper half plane.

Hence we can express them as $r_1 = re^{i\theta_1}$ and $r_2 = se^{i\theta_2}$, with $0 < \theta_1 \leq \theta_2 < \pi$.

Similarly, $r_1 r_2 = \frac{w}{3}$ tells us that $\theta_1 + \theta_2 = \angle w01 = \text{Arg}(w)$, the principal argument of w.

Consequently, we deduce that the angle between $\overline{0w}$ and $\overline{0r_2}$ equals θ_1. We note that, since $\mathcal{E}$ is tangent to the real axis at $x = \frac{1}{2}$, the origin 0 is external to $\mathcal{E}$.

Hence letting 0 play the role of A in lemma 3.2.7, we obtain that in fact $\mathcal{E}$ is tangent to the side $\overline{0w}$.

We now need to establish that the side from 1 to w is also tangent to $\mathcal{E}$. This is achieved by repeating the same proof, but with the triangle T reflected along the imaginary axis, that is, the vertices being -1, 0 and w. This would then complete the proof. □

We are now ready to give the proof of Marden's theorem as stated in 3.2.6. The lemmas discussed above greatly simplify our work and enable a short and concise proof:

Proof. (Marden's theorem). Assume $p(z)$, its roots and the triangle T are as in the statement of theorem 3.2.6.

We draw an ellipse $\mathcal{E}$ passing through the midpoint of one side of T, whose foci are at the critical points of $p(z)$. By lemma 3.2.9, $\mathcal{E}$ is tangent to this side of T. From lemma 3.2.10, we get that $\mathcal{E}$ is actually tangent to all three sides of T.

It remains to be shown that the points of tangency of the other two sides are at the midpoints. Suppose for the sake of contradiction that this is not the case. We then repeat the construction with a different side of the triangle T to produce another ellipse $\mathcal{E}'$.

But then, $\mathcal{E}$ and $\mathcal{E}'$ have the same foci, and are tangent to the same three sides, hence they coincide. Thus they are tangent to two sides of T at the midpoints. By symmetry, this applies to the remaining side of the triangle.

Thus the ellipse $\mathcal{E}$ is tangent to all sides of T at their midpoints, this completes the proof. □

Example 3.2.11. On the complex plane, we can choose non-collinear points r_1, r_2, r_3 as vertices of triangle $T = \triangle r_1 r_2 r_3$.

In this case:

$$\begin{aligned} p(z) &= (z - r_1)(z - r_2)(z - r_3) \\ &= z^3 - (r_1 + r_2 + r_3)z^2 + (r_1 r_2 + r_2 r_3 + r_1 r_3)z - r_1 r_2 r_3 \end{aligned}$$

hence:

$$p'(z) = 3\left(z^2 - \frac{2}{3}(r_1 + r_2 + r_3)z + \frac{1}{3}(r_1 r_2 + r_2 r_3 + r_1 r_3)\right)$$

Thus the roots of $p'(z)$ are the foci of the Steiner inellipse of T and are located at:

$$F_{\pm} = g \pm \sqrt{g^2 - \frac{1}{3}(z_1 z_2 + z_2 z_3 + z_1 z_3)}$$

where $g = \frac{1}{3}(z_1 + z_2 + z_3)$, as stated at the beginning of this section in Steiner's theorem.

In figure **??**, which we used to illustrate the Steiner inellipse, we chose the vertices of T to be $z_1 = 1+7i$, $z_2 = 7 + 5i$ and $z_3 = 3 + i$. This is the configuration for the polynomial $p(z) = (z - 1 - 7i)(z - 7 - 5i)(z - 3 - i)$. Constructing the corresponding Steiner inellipse, the foci are located at the points:

$$F_1 = \frac{1}{3}(11i + 13) \quad \text{and} \quad F_2 = 3 + 5i$$

These coincide with the roots of $p'(z)$, as predicted by Marden's theorem.

3.2.12 Some remarks on generalisations of Marden's theorem. Thus far, we have viewed Marden's theorem in light of cubic polynomials of the form $p(z) = (z - z_1)(z - z_2)(z - z_3)$, and consequently, in terms of a conic inscribed in a triangle. We note, however, that there are generalisations of this theorem to polynomials of higher degree, or equivalently, to $n - gons$ other than a triangle.

One such generalisation was studied by (Linfield, 1920) in 1920, who considered roots of rational functions, defined below. Before stating the result, we need the following definition.

Definition 3.2.13 (The class of a curve)**.** Let $\mathcal{C}$ be a curve and p a point generic to $\mathcal{C}$, that is, p does not lie on the curve. The class of $\mathcal{C}$ is the number of lines through p that are tangent to $\mathcal{C}$.

Remark 3.2.14. The class of a curve $\mathcal{C}$ is invariant with respect to the choice of p, as we count the number of tangent lines with multiplicities, including imaginary tangents.

We may now state Linfield's result, as considered in (Parish, 2006).

Theorem 3.2.15 (Linfield)**.** *Let $m_j \in \mathbb{R} \setminus \{0\}$, $j = 1, \ldots, n$ and let $z_1, \ldots, z_n$ be distinct complex numbers. Consider the rational function*

$$F(z) := \sum_{j=1}^{n} \frac{m_j}{z - z_j}$$

The zeros of $F(z)$ are the foci of the curve of class $n - 1$ which touches each line segment $\overline{z_i z_j}$ in a point dividing the line segment in the ratio $m_i : m_j$.

Remark 3.2.16. The function $F(z) = \sum_{j=1}^{n} \frac{m_j}{z-z_j}$ is obtained by taking the logarithmic derivative of $p(z) = \prod_{j=1}^{n}(z-z_j)^{m_j}$[3].

Hence, when $m_1, \ldots, m_n$ are positive integers, the zeros of $F(z)$ are the nontrivial critical points of $p(z)$. Marden's theorem then follows as a corollary when $n = 3$ and $m_1 = m_2 = m_3 = 1$.

Switching to the geometric side, we want to characterize polygons which admit an inscribed ellipse (or conic) which is tangent to all sides. These are often referred to as Steiner polygons. This approach brings us to the results found in (Argawal et al., 2013).

Theorem 3.2.17. *All convex $n-gons$ for $3 \leq n \leq 5$ are Steiner polygons, for $n \geq 6$, there exist $n-gons$ which are not Steiner polygons.*

Remark 3.2.18. Below we study a result from geometry that motivates the above generalisation of Marden's theorem.

Theorem 3.2.19 (Brianchon)**.** *A hexagon is a Steiner polygon if and only if its three diagonals meet at a point.*

Remark 3.2.20. A proof of the above result can be found in (The Art of Problem Solving). We note that, Brianchon's result only guarantees the existence of an inscribed ellipse, however, (Argawal et al., 2013) gives explicit constructions of these ellipses for $n - gons$ with $3 \leq n \leq 5$.

The reader should note that the case for $n \geq 6$ in the statement of theorem 3.2.17 can be deduced directly from Brianchon's result:

Given any three lines that cross each other pairwise, we can extend them to form diagonals of a hexagon. Amongst these lines, the case where all three meet at a point can be considered as the degenerate case and unlikely, thus, there must exist $n - gons$ which are not Steiner polygons when $n \geq 6$.

This brings us to the end of the study of the geometry of cubic polynomials.

[3]We utilize this technique in the study of Sendov's conjecture in the next chapter.

4. Sendov's Conjecture

We continue our study of the geometric relation between the roots and the critical points of a polynomial $p(z)$. Since the Gauss-Lucas theorem was proved, further attempts have been made to investigate this relation, the most significant of these results is Sendov's conjecture, which is an attempt to locate the critical points relative to each individual root of $p(z)$.

Before stating the conjecture, we begin with a useful proposition, whose proof can be found in (Sury, 2009).

4.1 Proving a special case of the conjecture

Proposition 4.1.1 (Getting a closed disc around the roots of $p(z)$)**.** Suppose $n \geq 1$, $a_0 \neq 0$ and let

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0.$$

Then the polynomial $z^n - |a_{n-1}|z^{n-1} - \cdots - |a_1|z - |a_0|$ has a unique, real and positive root M. Furthermore, $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$ has all its zeros in $D(0, M)$.

Remark 4.1.2. The above proposition thus enables us to study polynomials whose zeros lie in the unit disc without loss of generality. This is crucial in the statement of Sendov's conjecture.

We now state the conjecture in its general form:

Conjecture 4.1.3 (Sendov)**.** Let $p(z) = \prod_{j=1}^{n}(z - z_j)$ be a polynomial such that $|z_j| \leq 1$, $j = 1, \ldots, n$, that is, all its roots are inside the closed unit disc. Then each of the disks $|z - z_j| \leq 1$, $j = 1, \ldots, n$ contains a critical point of $p(z)$.

Several special cases of Sendov's conjecture have been verified, most notable of these is the paper by (Brown and Xiang, 1999), in which a proof of the conjecture for polynomials of degree $n \leq 8$, as well as arbitrary polynomials with at most 8 distinct zeros is presented.

Sendov's conjecture arose as a consequence of studying the polynomial $p(z) = z^n - 1$ together with its rotations, that is, $p(z) = z^n - e^{i\theta}$, $\theta \in \mathbb{R}$. These are called the extremal polynomials. We recall, from definition 2.2.7 in chapter two, for $Z(p)$ being the set of roots of $p(z)$, we defined:

$$T(p) = \max\{|w - v| : w \in Z(p), v \in Z(p')\ w \neq v\}.$$

With this in mind, the Gauss-Lucas theorem implies that $T(p) \leq 2$ for $p(z)$ as defined in the statement of the conjecture. Hence, we may consider Sendov's conjecture as a refinement of the Gauss-Lucas theorem, since if the conjecture is true, then this would imply that actually $T(p) \leq 1$.

In the paper by (Bojanov et al., 1985), an improvement was made on the bound of $T(p)$, and the authors showed that $T(p) \leq 1.0833164$. Furthermore, it was shown that:

$$T(p) \longrightarrow 1 \quad \text{as} \quad n \longrightarrow \infty.$$

In this section, we state and prove the conjecture for polynomials with zeros on the unit circle, this version of the proof is due to (Sheil-Small, 2002):

Theorem 4.1.4. *(Sheil-Small, 2002) Let*

$$p(z) = \prod_{j=1}^{n}(z - z_j) \quad \textit{where} \quad |z_j| \leq 1, \quad j = 1, \dots, n.$$

Pick a root r of $p(z)$, where $|r| = 1$. Then $p'(z)$ has a root in the disc:

$$\left\{ z \in \mathbb{C} : \left| z - \frac{r}{2} \right| \leq \frac{1}{2} \right\}$$

Proof. We can write the polynomial $p(z)$ as:

$$p(z) = (z - r)\prod_{j=1}^{n-1}(z - z_j) = (z - r)S(z)$$

and note that the zeros of $S(z)$, $z_1, \dots, z_{n-1}$ are also in the unit disc.

We assume furthermore that all the zeros of $S(z)$ are distinct from r, as the result would follow trivially by multiplicity of r.

Let $\zeta_1, \zeta_2, \dots, \zeta_{n-1}$ be the critical points of $p(z)$. We note:

$$\begin{aligned} p(z) &= (z - r)S(z) \\ \implies p'(z) &= (z - r)S'(z) + S(z) \\ \implies p''(z) &= (z - r)S''(z) + 2S'(z) \end{aligned}$$

Therefore:

$$\frac{p''(z)}{p'(z)} = \frac{(z - r)S''(z) + 2S'(z)}{(z - r)S'(z) + S(z)}$$

This implies that:

$$\frac{p''(r)}{p'(r)} = 2\frac{S'(r)}{S(r)}$$

This is equivalent to taking the derivative on both sides of the equality:

$$\log(p'(z)) = 2\log(S(z)),$$

at the point $z = r$, thus we obtain:

$$\sum_{j=1}^{n-1} \frac{r}{r-\zeta_j} = 2\sum_{j=1}^{n-1} \frac{r}{r-z_j}$$

Therefore:

$$Re\left(\sum_{j=1}^{n-1} \frac{r}{r-\zeta_j}\right) = 2Re\left(\sum_{j=1}^{n-1} \frac{r}{r-z_j}\right) = 2Re\left(\sum_{j=1}^{n-1} \frac{1}{1-z_j\bar{r}}\right) \geq n-1 \tag{4.1.1}$$

Where the last inequality is due to the Gauss-Lucas theorem and proposition 3.1.12. That is:

$$\text{each} \quad |z_j\bar{r}| \leq 1, \quad \text{hence} \quad Re\left(\frac{1}{1-z_j\bar{r}}\right) \geq \frac{1}{2}, \quad j = 1,\ldots,n-1.$$

Thus from equation 4.1.1, it follows that for at least one j, we have:

$$Re\left(\frac{r}{r-\zeta_j}\right) \geq 1.$$

Equivalently, we have that:

$$\left|\zeta_j - \frac{r}{2}\right| \leq \frac{1}{2}$$

and this completes the proof. □

Remark 4.1.5. Upon closer scrutiny of our proof, we note that it cannot be applied in general, when $0 < |r| < 1$, since any z_j such that $|z_j| > |r|$ will contribute negatively to the right hand side of equation 4.1.1, eventually contradicting the inequality. The above proof was motivated by attempting to prove the conjecture for polynomials $p(z) \in \Gamma$ as defined in the previous chapter on the geometry of cubic polynomials.

On the other hand, although the approach in the proof is not sufficient to prove the general conjecture, it has motivated the discussion of the following section, where we suggest an approach that could be considered towards resolving the conjecture.

4.2 A possible approach towards resolving the conjecture

We begin this section by proving a general result about centroids of polynomials:

Definition 4.2.1 (centroid of a polynomial)**.** For the polynomial $p(z) = a_nz^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ with n distinct roots r_j, $j = 1,\ldots,n$, we define the centroid of $p(z)$ to be the arithmetic mean of the zeros of $p(z)$, that is:

$$C_p := \frac{1}{n}\sum_{j=1}^{n} r_j = \frac{r_1 + r_2 + \cdots + r_n}{n}.$$

Proposition 4.2.2. The centroid of a polynomial is invariant under differentiation.

Proof. It suffices to show that $C_p = C_{p'}$. Since:

$$p(z) = \sum_{k=1}^{n} a_k z^k = a_n \prod_{k=1}^{n} (z - r_k),$$

we proceed by comparing coefficients. Expanding the first equality we get:

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \tag{4.2.1}$$

and from the second equality:

$$p(z) = a_n[(z - r_1)(z - r_2) \cdots (z - r_n)].$$

Expanding the product in the square brackets gives:

$$\begin{aligned} p(z) &= a_n[z^n - (r_1 + r_2 + \cdots + r_n)z^{n-1} + \cdots + (-1)^n r_1 r_2 \cdots r_n & (4.2.2) \\ &= a_n z^n - a_n(r_1 + r_2 + \cdots + r_n)z^{n-1} + \cdots + (-1)^n a_n r_1 r_2 \cdots r_n & (4.2.3) \end{aligned}$$

Comparing the coefficients of z^{n-1} from equations 4.2.1 and 4.2.3 we get:

$$a_{n-1} = -a_n(r_1 + \cdots + r_n)$$

and hence the centroid of $p(z)$ is given by the identity:

$$\frac{1}{n}(r_1 + r_2 + \cdots + r_n) = -\frac{a_{n-1}}{n a_n}. \tag{4.2.4}$$

On the other hand:

$$p'(z) = \sum_{k=0}^{n-1} (k+1) a_{k+1} z^k = n a_n \prod_{k=1}^{n-1} (z - \zeta_k)$$

Hence, by the identity from equation 4.2.4, the centroid of $p'(z)$ is:

$$C_{p'} = \frac{1}{n-1}(\zeta_1 + \cdots + \zeta_{n-1}) = -\frac{1}{n-1}\left(\frac{(n-1)a_{n-1}}{n a_n}\right) = -\frac{1}{n}\frac{a_{n-1}}{a_n} = C_p.$$

This concludes the proof of the proposition. ∎

Our method is an attempt to find a polynomial $p(z)$ that would give a counterexample to Sendov's conjecture, or equivalently, show that if such a polynomial exists, it would violate some given conditions that should normally be satisfied.

We can articulate these ideas better through the example below.
Example 4.2.3. Suppose the polynomial $p(z)$ has all its zeros in the unit disc. We then pick r, a particular root of $p(z)$, hence we have that:

$$r \in \{z \in \mathbb{C} : |z| \leq 1\} =: \mathcal{Z}$$

According to Sendov's conjecture, there is a critical point c satisfying:

$$c \in \{z \in \mathbb{C} : |z - r| \leq 1\} =: \mathcal{C}$$

Therefore $c \in \mathcal{C} \cap \mathcal{Z}$ by the Gauss-Lucas theorem. Hence for $p(z)$ to be a counterexample to the conjecture, all critical points of $p(z)$ should satisfy:

$$c \in \mathcal{Z} \setminus \mathcal{C} \cap \mathcal{Z}.$$

In particular, we consider the case where the centroid C_p is also in $\mathcal{Z} \setminus \mathcal{C} \cap \mathcal{Z}$.

This has profound implications on the form of the polynomial $p(z)$.

For instance, from proposition 4.2.2 and the location of C_p, what can we deduce about the distribution of the zeros of $p(z)$ around the unit disc?

Consequently, could we find a polynomial $p(z)$ in the unit disc satisfying such a distribution of zeros?

On the other hand, exactly how far can a root r or a critical point c be from the centroid?

We believe answers to these questions would shed more light, not only on the conjecture, but will also further enhance our understanding of the geometry of polynomials.

Remark 4.2.4. As a preliminary result, let us define the minimum distance from the centroid to the nearest critical point of $p(z)$, of degree $n > 1$:

$$\alpha_n := \min\{|C_p - \zeta_k| : \zeta_k \in Z(p')\}.$$

For $p(z) = z^n - z$, all but one zero of $p(z)$ lie on the unit circle, hence we regard it as an extremal polynomial, we note that:

$$p'(z) = nz^{n-1} - 1,$$

hence the critical points of $p(z)$ satisfy $z^{n-1} = \frac{1}{n}$, therefore:

$$|\zeta_k| = \frac{1}{n^{\frac{1}{n-1}}},$$

a lower bound for α_n.

On the other hand, by proposition 4.2.2, the centroids of $p(z)$ (not necessarily $z^n - z$) and $p'(z)$ coincide, and by the Gauss-Lucas theorem, the roots of $p'(z)$ are in the unit disc, hence at least one of them is within a unit distance from the centroid. Hence for $p(z)$ with zeros in the unit disc:

$$\frac{1}{n^{\frac{1}{n-1}}} \leq \alpha_n \leq 1.$$

We believe that these bounds can be sharpened further, and together with other results relating critical points, roots and the centroid of a polynomial, could be used to construct a counterexample to Sendov's conjecture or to prove the non existence of such a polynomial, in either case resolving the conjecture.

We conclude this section with a figure illustrating the geometric configuration of the idea from example 4.2.3.

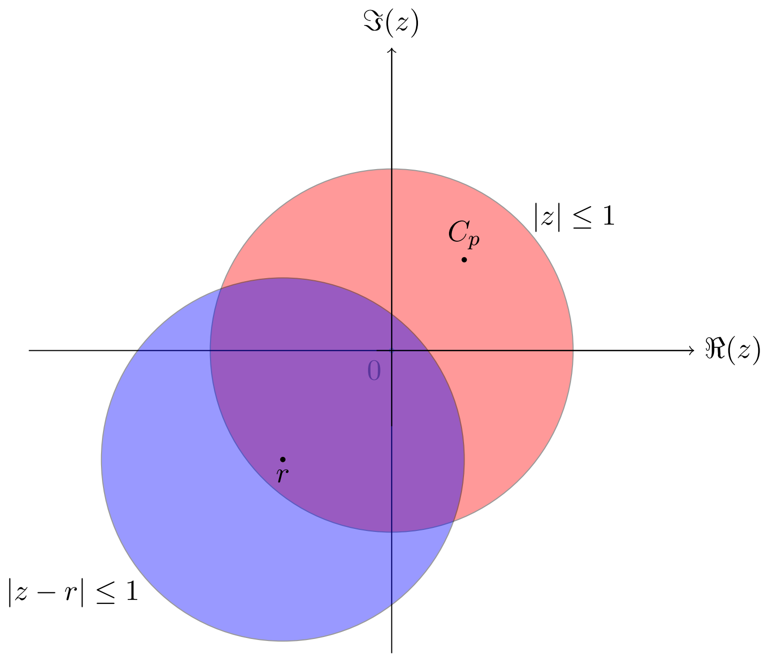

Figure 4.1: Illustrating the location of C_p relative to r.

5. Conclusion

We began this essay with a proof of the Gauss-Lucas theorem, which states that the critical points of a complex polynomial $p(z)$ lie in the convex hull of its roots. We then showed, through a refinement of the theorem, that actually the critical points lie inside a smaller region strictly contained in the boundary of the convex hull.

We then specialised our study to cubic polynomials, first characterising the polynomials by the location of their critical points, and hence leading to the discussion and proof of Marden's theorem. We saw that, for a cubic polynomial $p(z)$ with three distinct zeros z_1, z_2 and z_3, the Gauss-Lucas theorem tells us that the roots of $p'(z)$ lie in the triangle $\triangle z_1 z_2 z_3$. Marden's theorem, however, gives a sharper prediction for the location of these critical points, characterising them as the foci of the Steiner inellipse inscribed in triangle $\triangle z_1 z_2 z_3$.

Finally, we discussed Sendov's conjecture, giving a proof of a special case, and motivated by the results from previous sections, we suggested a possible method which could be considered towards resolving the conjecture.

However, we admit that we should still be sceptical about this approach towards the problem, as more preliminary investigations have to be carried out to determine if this path would yield any more meaningful insights into the conjecture.

Acknowledgements

To Florian, Barry, Jeff, Jan, Frances, Igsaan, Danny, Talin and the entire AIMS 2013-14 family, thank you!

References

M. Argawal, J. Clifford, and M. Lachance. Implications of marden's theorem for inscribed ellipses. *American Mathematical Monthly*, 121(1), October 2013.

B. D. Bojanov, Q. I. Rahman, and J. Szynal. On a conjecture of sendov about the critical points of a polynomial. *Mathematische Zeitschrift*, 190(2):281–285, January 1985.

J. E. Brown and G. Xiang. Proof of the sendov conjecture for polynomials of degree at most eight. *Journal of Mathematical Analysis and Applications*, 232(2):272–292, April 1999.

B. Curgus and V. Mascioni. On the location of critical points of polynomials. *Proceedings of the American Mathematical Society*, 131(1):253–264, June 2003.

B. Curgus and V. Mascioni. A contraction of the lucas polygon. *Proceedings of the American Mathematical Society*, 132(10):2973–2981, May 2004.

C. Frayer, M. Kwon, C. Schafhauser, and J. A. Swenson. The geometry of cubic polynomials. *Mathematics Magazine*, 87(2):113–124, April 2014.

W. K. Hayman. *Research Problems in Function Theory*. London: Athlone Press, 1967.

D. Kalman. An elementary proof of marden's theorem. *American Mathematical Monthly*, 115(4): 330–338, 2008.

B.-Z. Linfield. On the relation of the roots and poles of a rational function to the roots of its derivative. *Bulletin of the American Mathematical Society*, 27(1):17–21, 1920.

M. Marden. *Geometry of Polynomials*. American Mathematical Society, 1966.

D. Minda and S. Phelps. Triangles, ellipses and cubic polynomials. *American Mathematical Monthly*, 115(8):679–689, October 2008.

T. Needham. *Visual Complex Analysis*. Oxford University Press, 1999.

J. L. Parish. On the derivative of a vertex polynomial. *Forum Geometricorum*, 6(•):285–288, 2006.

T. Sheil-Small. *Complex Polynomials*. Cambridge University Press, 2002.

B. Sury. When are complex zeroes on a roll(e)? *Resonance*, 14(9):872–881, September 2009.

The Art of Problem Solving. Brianchon's theorem. The Art of Problem Solving, http://www.artofproblemsolving.com/blog/38117, Accessed April 2014.

Wikipedia. Convex hull. Wikipedia, the Free Encyclopedia, http://en.wikipedia.org/wiki/Convex_hull, Accessed April 2014a.

Wikipedia. Marden's theorem. Wikipedia, the Free Encyclopedia, http://en.wikipedia.org/wiki/Marden's_Theorem, Accessed April 2014b.

Wolfram Mathworld. Rolle's theorem. Wolfram Mathworld, http://mathworld.wolfram.com/RollesTheorem.html, Accessed April 2014.