McNair Scholars Journal

Volume 13 | Issue 1 Article 3

2009

Investigations in the Geometry of Polynomials

Neil Biegalle
Grand Valley State University

Investigations in the Geometry of Polynomials

Neil Biegalle
McNair Scholar

Matt Boelkins, Ph.D.
Faculty Mentor

Abstract

Because polynomial functions are completely determined by their roots, every property of a polynomial is affected when these roots change. Our research aims to further our understanding of how the distribution of a polynomial's roots affects specjfic characteristics of the function. We are especially interested in classifying which root distributions maximize or minimize certain properties. We employ recent results on polynomial root dragging and root motion to explore these issues further, including the attempt to explain why many properties are maximized by Bernstein polynomials. This paper will survey some important results and present our investigations into new problems and approaches.

1 Introduction

How do changes in the roots of a monic polynomial with all real zeros affect the other characteristics of the function? This is the question that drives our investigations in the geometry of polynomials. It is well known that a monic polynomial with all real zeros is uniquely determined by the placement of these zeros. That is, given any n roots, there is exactly one monic polynomial with all real zeros that passes through these roots. Therefore, every characteristic of a polynomial depends on the location of its roots, and even the slightest change in the root distribution will produce an entirely new polynomial. Many questions in this same field of study have been answered, and the beginning section of this paper will be focused on introducing some of these important results, along with the key concepts that were necessary for proving them. The goal of this paper is to contribute to the overall understanding that we possess of this relationship between a polynomial's roots and its other characteristics.

For the sake of convenience, and since any polynomial with all real zeros may be scaled to be monic with all roots in $[-1, 1]$, we have narrowed the focus of this paper to monic polynomials with all real zeros that live in the interval $[-1, 1]$. That is, we are interested in polynomials that can be written in the form

$$p(x) = \prod_{i=1}^{n}(x - r_i),$$

where n is the degree of the polynomial, and the r_i's are real numbers in $[-1, 1]$ that represent the roots. This provides us with some consistency and makes it sensible to compare properties from one polynomial to another.

Once we are able to compare polynomials and the properties that they possess, we may then ask a very natural question that is so common in mathematics: "What root location will make property X the most extreme?" This question of extremality has been the primary focus of our research, and it is in this spirit that we write this paper. There has been much work done in this field in recent years, and the results developed provide crucial perspectives and tools for our investigations. The Polynomial Root Dragging Theorem [1, 2] is one of the most influential theorems in our research, and we will illustrate how extremal problems in the geometry of polynomials may be thought of more intuitively from this perspective.

One intriguing question that has been given little attention arises when we notice the pronounced patterns found when analyzing problems dealing with maximality. A large portion of this paper will be dedicated to surveying some of the results that are available pertaining to these maximal problems. While the individual problems that we will outline have well-established proofs, we strive to offer what we believe to be a promising, yet undeveloped, general explanation for the "maximal polynomial" phenomenon. Along the way, we make some additional observations that are either new or underreported.

2 Fundamental Results

First and foremost, it is important to survey some fundamental theorems and concepts in the geometry of polynomials. This section is meant to clarify the sort of problems that this research is interested in answering, as well as give the reader a flavor for the different kinds of mathematics that are employed when investigating this kind of problem. Here, we outline some of these important theorems and extend a couple of them to prove new results. The results themselves, as well as the analytic approach used, will be highlighted, and important concepts will be defined appropriately.

As mentioned earlier, one of the most important results in this field is the Polynomial Root Dragging Theorem [1, 2], as proven by Bruce Anderson. This theorem illustrates the effect that "dragging" any number of roots of a polynomial in a given direction has on the location of the critical numbers of the polynomial. The proof uses a concept known as the logarithmic derivative. As this concept appears frequently in the geometry of polynomials, and specifically in the proofs we outline in this paper, it is worth discussing in some detail before we present the theorem.

Definition 1. *The logarithmic derivative of a function $p(x)$ is the quantity*

$$\frac{p'(x)}{p(x)}.$$

This is called the logarithmic derivative because it is obtained by taking the derivative of the logarithm of a given polynomial p: $\frac{d}{dx}[\ln(p(x))] = \frac{p'(x)}{p(x)}$. For a given degree n monic polynomial $p(x) = \prod_{i=1}^{n}(x - r_i)$, with roots $r_1, r_2, \ldots, r_n$, the logarithmic derivative is

$$\frac{p'(x)}{p(x)} = \sum_{i=1}^{n} \frac{1}{(x - r_i)}.$$

This is a strictly decreasing, rational function with $n-1$ zeros, and it proves to be a very useful tool in multiple theorems throughout this paper.

Now we may introduce the Polynomial Root Dragging Theorem. As we will show, in response to dragging a polynomial's root(s) in a given direction, its critical numbers will move in the same direction as the roots that are dragged, or they will remain in place. Further, the critical numbers that move will move less than the root that is moved. Here, we prove this result for moving roots to the right. The proof can be easily altered to prove that the result is also true when we move roots to the left.

Theorem 1. *Let $p(x)$ be a polynomial of degree n with distinct real roots $r_1, r_2, \ldots, r_n$. Then as we "drag" some or all of the interior roots a distance at most ε to the right, the critical points will all follow to the right, and each of them will move less than ε units.*

Proof. Let $p(x)$ be a polynomial of degree n with distinct real roots $r_1, r_2, \ldots, r_n$. We will prove that as we "drag" some or all of the interior roots a distance at most ε to the right, the critical points will all follow to the right, and each of them will move less than ε units. Letting $p(x) = (x - r_1)(x - r_2)\cdots(x - r_n)$, we know that for any critical point c, $p'(c) = 0$. Since there are no repeated roots, we know that $p(c) \neq 0$. Therefore, we can take the opposite of the logarithmic derivative and obtain

$$-\frac{p'(c)}{p(c)} = \sum_{i=1}^{n} \frac{1}{r_i - c} = 0. \tag{1}$$

So equation (1) shows that c is an implicit function of the roots of $p(x)$. Implicit differentiation with respect to r_i gives

$$\frac{\partial}{\partial r_i}\left(\frac{1}{r_1 - c}\right) + \frac{\partial}{\partial r_i}\left(\frac{1}{r_2 - c}\right) + \cdots + \frac{\partial}{\partial r_i}\left(\frac{1}{r_i - c}\right) + \cdots + \frac{\partial}{\partial r_i}\left(\frac{1}{r_n - c}\right) = \frac{\partial}{\partial r_i}(0). \tag{2}$$

This yields

$$\left(\frac{1}{r_1 - c}\right)^2 \frac{\partial c}{\partial r_i} + \cdots + \left(\frac{1}{r_{i-1} - c}\right)^2 \frac{\partial c}{\partial r_i} + \left[\left(\frac{1}{r_i - c}\right)^2 \frac{\partial c}{\partial r_i} - \left(\frac{1}{r_i - c}\right)^2\right] + \left(\frac{1}{r_{i+1} - c}\right)^2 \frac{\partial c}{\partial r_i} + \cdots + \left(\frac{1}{r_n - c}\right)^2 \frac{\partial c}{\partial r_i} = 0. \tag{3}$$

From here, solving for $\frac{\partial c}{\partial r_i}$ leads us to see that

$$\frac{\partial c}{\partial r_i} = \frac{\left(\frac{1}{r_i - c}\right)^2}{\left(\frac{1}{r_1 - c}\right)^2 + \left(\frac{1}{r_2 - c}\right)^2 + \cdots + \left(\frac{1}{r_n - c}\right)^2}.$$

Clearly, $\frac{\partial c}{\partial r_i}$ is positive. Further, if we let $A \subset \{1, 2, \ldots, n\}$ with $|A| < n$ it follows that

$$0 < \sum_{i \in A} \frac{\partial c}{\partial r_i} < 1. \tag{4}$$

Since c depends on the roots, we can now consider what happens to c when we shift each of the interior roots r_i to the right by $\varepsilon_i \geq 0$, respectively. We can define a function f that represents the location of c after the roots have been moved as follows:

$$f(t) := c(r_1, r_2 + t\varepsilon_2, \ldots, r_{n-1} + t\varepsilon_{n-1}, r_n)$$

where t varies from 0 to 1. Then by the multivariable chain rule, we have

$$f'(t) = \sum_{2}^{n-1} \frac{\partial c}{\partial r_i}\frac{dr_i}{dt} = \sum_{2}^{n-1} \frac{\partial c}{\partial r_i}\varepsilon_i \leq \max(\varepsilon_i) \sum_{2}^{n-1} \frac{\partial c}{\partial r_i}.$$

By (4), we can conclude that

$$\max(\varepsilon_i) \sum_{2}^{n-1} \frac{\partial c}{\partial r_i} < \max(\varepsilon_i),$$

and thus $f'(t) < \max(\varepsilon_i)$. The mean value theorem states that for some ζ between 0 and 1, $f'(\zeta) = f(1) - f(0)$. Since we have shown that for all t between 0 and

1, $f'(t) < \max(\varepsilon_i)$, we can then conclude that $f(1) - f(0) < \max(\varepsilon_i)$. Therefore, c will move to the right strictly less than the root that moves the most.

$\square$

This result has been especially influential in our research, primarily because it offers such a unique and valuable approach to analyzing the effects that changing roots has on a polynomial. This approach is an important tool that will help us in furthering our understanding of how roots and other characteristics of a polynomial interact.

When this theorem is first encountered, it is very natural to wonder exactly how far the critical numbers are moving. Following, we address the question of how far the average critical number travels in response to dragging a root ε units. However, to do so, we must first introduce the concept of the *centroid* of a polynomial $p(x)$, which is denoted by A_p.

Definition 2. *Given a polynomial $p(x)$, the centroid of p, or A_p is the average of the roots of $p(x)$. That is, if $p(x)$ has roots $r_1, r_2, \ldots, r_n$, then*

$$A_p = \frac{r_1 + r_2 + \cdots + r_n}{n}.$$

In 1998, Piotr Pawlowki [11] noted that the centroid of a polynomial is differentiation invariant. In other words, for any given polynomial, the average root is the same as the average critical number, which is the same as the average inflection point, and so on. This is an amazing feature of polynomials, and it allows us to quantify the average distance a critical number moves in reaction to dragging a root ε units. We will establish this result with a short corollary.

Corollary 1. *Let p be a monic, degree n polynomial with all real zeros $r_1, r_2, \ldots r_n$. If we drag a single root, r_i, ε units to the right, the critical numbers of p will move to the right by an average of $\frac{\varepsilon}{n}$ units.*

Proof. Let p be a monic, degree n polynomial with all real zeros $r_1, r_2, \ldots r_n$. Then the centroid of p is given by

$$A_p = \frac{r_1 + r_2 + \cdots + r_n}{n}.$$

If we drag the root r_i by ε units to the right, we will then have a new polynomial, p_ε, whose centroid is

$$\begin{aligned} A_{p_\varepsilon} &= \frac{r_1 + r_2 + \cdots + (r_i + \varepsilon) + \cdots + r_n}{n} \\ &= \frac{r_1 + r_2 + \cdots + r_n}{n} + \frac{\varepsilon}{n} \\ &= A_p + \frac{\varepsilon}{n}. \end{aligned}$$

Since the centroid is differentiation invariant, we have

$$A_{p'_\varepsilon} = A_{p'} + \frac{\varepsilon}{n}.$$

Therefore, moving r_i to the right ε units has caused the average critical number of p to move to the right $\frac{\varepsilon}{n}$ units. $\square$

Since there are $n-1$ critical numbers for a degree n polynomial, it may also be noted that the total sum of the distance travelled by the critical numbers in response to a single root being dragged ε units is $\frac{(n-1)}{n}\varepsilon$. So when dragging a single root, not only does each critical number move less than the root that is dragged, but all of the critical numbers combined move less than the root that is dragged. We may further consider what happens if we allow for the dragging of m roots, where $m \leq n$, and the m roots move distances of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$, respectively. It is simple to show that the average critical number of the polynomial would then move

$$\frac{\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_m}{n}.$$

Note that this last result is true regardless of the sign of the $\varepsilon_i's$.

This idea of allowing roots to be dragged in opposite directions creates another question that the Polynomial Root Dragging Theorem does not answer. What happens to critical numbers of a polynomial in reaction to dragging roots in opposite directions? There is a theorem that is closely related to the Polynomial Root Dragging Theorem, called the Polynomial Root Squeezing Theorem [4], which addresses exactly this. This theorem may be thought of as an extension of the Root Dragging Theorem, and it has some interesting consequences considering the span of the derivative of a polynomial. For future reference, we will define here what is meant by *span*.

Definition 3. *The span of a polynomial is the distance between the least and greatest roots of the polynomial.*

Now, we will introduce the Polynomial Root Squeezing Theorem, followed by its proof.

Theorem 2. *Let p be a monic, degree n polynomial with all real roots $r_1 \leq r_2 \leq \cdots \leq r_n$ such that $r_1 = -b$ and $r_n = b$ for some positive real number b. Further, let $c_1 \leq c_2 \leq \cdots \leq c_{n-1}$ be the critical numbers of p. Let $r_j < r_k$ be any two interior roots of p and $d \in \mathbb{R}^+$ be such that*

$$d \leq \min\Big\{r_{j+1} - r_j, r_k - r_{k-1}, \frac{1}{2}(r_k - r_j)\Big\}.$$

Let $\tilde{p}$ be the polynomial that results from squeezing r_j and r_k together by a distance $2d$. That is,

$$\tilde{p}(x) = (x - r_j - d)(x - r_k + d) \prod_{i=1, i\neq j,k}^{n} (x - r_i).$$

Denote the critical points of $\tilde{p}$ by $\tilde{c}_1 \leq \tilde{c}_2 \leq \cdots \leq \tilde{c}_{n-1}$. Then for $1 \leq i < j$ we have $\tilde{c}_i \geq c_i$, and for $k \leq i \leq n-1$ we have $\tilde{c}_i \leq c_i$.

Proof. Let p be a monic, degree n polynomial with all real roots $r_1 \leq r_2 \leq \cdots \leq r_n$ such that $r_1 = -b$ and $r_n = b$ for some positive real number b. Let r_j and r_k be interior roots of p such that $r_j < r_k$. Further, let c_i be a critical number of p such that $1 \leq i < j$ or $k \leq i \leq n-1$. Let $d \in \mathbb{R}^+$ such that

$$d \leq \min\left\{r_{j+1} - r_j, r_k - r_{k-1}, \frac{1}{2}(r_k - r_j)\right\}.$$

Define $\tilde{p}$ to be the polynomial that results from squeezing r_j and r_k together by a distance $2d$, with critical points $\tilde{c}_1 \leq \tilde{c}_2 \leq \cdots \leq \tilde{c}_{n-1}$. That is,

$$\tilde{p}(x) = (x - r_j - d)(x - r_k + d) \prod_{i=1, i\neq j,k}^{n} (x - r_i).$$

If c_i is a root of the polynomial, then that root is repeated and we have three possible cases:

1. $r_i = c_i = r_{i+1}$ and neither r_i nor r_{i+1} are being shifted.
2. $r_i = c_i = r_{i+1}$ and r_{i+1} is being shifted to the right.
3. $r_i = c_i = r_{i+1}$ and r_i is being shifted to the left.

In case 1, since $\tilde{c}_i$ must still be the critical number between r_i and r_{i+1}, we have $r_i = \tilde{c}_i = r_{i+1} = c_i$, as desired. In case 2, according to Rolle's Theorem, $\tilde{c}_i$ must be greater than r_i. Therefore we have $\tilde{c}_i > c_i$. This is what we desired to show, as the fact that r_{i+1} is moving to the right implies that $1 \leq i < j$. Similarly, in case 3, Rolle's Theorem tells us that $\tilde{c}_i$ must be less than r_{i+1}, and hence, less than c_i. This is what we desired to show, as the fact that r_i is moving to the left implies that $k \leq i \leq n-1$.

We must now investigate what happens in the case where $r_i < c_i < r_{i+1}$. Since our goal is to develop a way to compare c_i and $\tilde{c}_i$, we can examine how $\tilde{p}'$ behaves near these points. We define the function $q(x)$ so that $p(x) = (x - r_j)(x - r_k)q(x)$. From this, we can rewrite $\tilde{p}(x)$ as $\tilde{p}(x) = (x - r_j - d)(x - r_k + d)q(x)$. Differentiating $p(x)$ yields

$$p'(x) = (x - r_j + x - r_k)q(x) + (x - r_j)(x - r_k)q'(x). \quad (5)$$

Differentiating $\tilde{p}(x)$ yields

$$\begin{aligned}\tilde{p}'(x) = (x - r_j + x - r_k)q(x) \\ + (x - r_j - d)(x - r_k + d)q'(x). \quad (6)\end{aligned}$$

We may now subtract equation (5) from equation (6) and obtain

$$\begin{aligned}
\tilde{p}'(x) - p'(x) \quad &= [(x - r_j + x - r_k)q(x) \\
&\quad + (x - r_j - d)(x - r_k + d)q'(x)] \\
&\quad - [(x - r_j + x - r_k)q(x) \\
&\quad + (x - r_j)(x - r_k)q'(x)] \\
&= (x - r_j - d)(x - r_k + d)q'(x) \\
&\quad - (x - r_j)(x - r_k)q'(x) \\
&= [(x^2 - xr_k + dx - r_jx + r_jr_k \\
&\quad - r_jd - dx + dr_k - d^2) \\
&\quad - (x^2 - xr_k - r_jx \\
&\quad + r_jr_k)]q'(x) \\
&= (-r_jd + dr_k - d^2)q'(x) \\
&= d(r_k - r_j - d)q'(x).
\end{aligned}$$

Since $p'(c_i) = 0$, we can then evaluate this expression at $x = c_i$ to find that

$$\tilde{p}'(c_i) = d(r_k - r_j - d)q'(c_i). \quad (7)$$

Recall our assumption that $(r_k - r_j) \geq 2d$. Therefore, we have $(r_k - r_j - d) \geq d$, which we have assumed to be positive. Therefore, $\tilde{p}'(c_i)$ must have the same sign as $q'(c_i)$.

Now, consider the case where $c_i < r_j$. Clearly, for $r_i < x < r_{i+1}$, $p(x)$ is either strictly positive or negative. Let us assume that $p(x)$ is negative. We then know that $p(c_i)$ is negative. Since $p(c_i) = (c_i - r_j)(c_i - r_k)q(c_i)$, and since $(c_i - r_j)$ and $(c_i - r_k)$ are both negative, we can see that $q(c_i)$ must also be negative. If we now evaluate equation (5) at $x = c_i$, we obtain

$$0 = (c_i - r_j + c_i - r_k)q(c_i) + (c_i - r_j)(c_i - r_k)q'(c_i). \quad (8)$$

We know that $(c_i - r_j + c_i - r_k)q(c_i)$ is positive, since both factors are negative. Similarly, we know that $(c_i - r_j)(c_i - r_k)$ is positive. Therefore, to satisfy equation (8), we see that $q'(c_i)$ must be negative. Further, by equation (6), this tells us that $p'(c_i)$ must be negative as well. We may now consider the equation

$$p(x)(x - r_j - d)(x - r_k + d) = \tilde{p}(x)(x - r_j)(x - r_k). \quad (9)$$

Evaluating equation (9) at $x = c_i$, the left hand side of the equation becomes negative based on what we have shown so far. Therefore, $\tilde{p}(c_i)$ must be negative to make the equation true.

Let us refer to the ith root of $\tilde{p}(x)$ as $\tilde{r}_i$. Since we are still under the assumption that $i < j$, we know that $\tilde{r}_i = r_i$ and $r_{i+1} \leq \tilde{r}_{i+1}$. Therefore, when x is in the interval (r_i, r_{i+1}), x is also in the interval $(\tilde{r}_i, \tilde{r}_{i+1})$. Again, since the function $\tilde{p}(x)$ must be strictly positive or strictly negative over this interval, the fact that $\tilde{p}(c_i)$ is negative lets us conclude that $\tilde{p}(x)$ is negative on the entire interval $(\tilde{r}_i, \tilde{r}_{i+1})$. So we know that the sign of $\tilde{p}'(x)$ must change from negative to positive exactly once, namely at $\tilde{c}_i$, in this interval. We have shown that $\tilde{p}'(c_i)$ must be negative, and therefore, c_i must be less than $\tilde{c}_i$, which is what we set out to prove.

This argument is similar if we assume that $p(x)$ is positive for $r_i < x < r_{i+1}$.

Further, the same reasoning may be applied to show that for $k \leq i \leq n-1$ we have $\tilde{c}_i \leq c_i$. This completes our proof.

□

A recent paper by Christopher Frayer and James Swenson [7] takes a more dynamic approach to this same problem of how critical numbers move with respect to changing roots. In their paper, entitled Continuous Polynomial Root Dragging, they let each root move with a prescribed velocity, and the location of the roots is then thought of as a function of time. They then show, at a given time, which way a specified critical point is moving. They call it the Polynomial Root Motion Theorem, and it is stated as follows:

Theorem 3. *Suppose that $c(t)$ is a critical point of the polynomial $p_t(x)$ for all t, with c differentiable and $c(t_0) = 0$. Further, for each root r_k, the root is moving at a velocity of v_k. If $p_{t_0}(x)$ has a double root at $x = 0$ (say, $r_k = r_k(t_0) = 0$ for $k \in \{i, j\}$), then $c'(t_0) = (v_i + v_j)/2$. Otherwise,*

$$c'(t_0) = \frac{-p_{t_0}(0)}{p''_{t_0}(0)} \sum_{i=1}^{n} \frac{v_i}{r_i^2}.$$

This theorem shows that roots affect critical points in almost the same way that gravity affects masses. Note that it assumes that the critical point of interest is at the origin. However, since any polynomial may be translated horizontally to make a given critical point land on the origin, this theorem can be used to show how any given critical point is moving at a given time.

We now have a much better understanding of how critical numbers interact with changing roots. However, the critical numbers are not the only things that change when we change roots. Remember, *every* characteristic is dependent upon the location of the roots. So how do other things change? The following result arises as another consequence of root dragging, and it is another great example of the kind of things we are trying to analyze. It appeared in a paper by Matt Boelkins, Jennifer Miller, and Benjamin Vugteveen in 2006 [5]. It uses a similar analysis as the Polynomial Root Dragging Theorem to illustrate the effect that dragging roots has on the curve's deviation from the x-axis. This theorem will be used later when we begin to explore some extremal problems.

Theorem 4. *Let $p(x) = \prod_{i=1}^{n}(x - r_i)$ be a degree n polynomial where $r_1 < r_2 < \cdots < r_n$. If an interior root, r_d, is dragged to the right by a distance of ε, where $0 < \varepsilon < r_{d+1} - r_d$, producing a new polynomial, $p_\varepsilon(x) = (x - r_d - \varepsilon)\prod_{i \neq d}(x - r_i)$, then the following inequalities hold:*

a. If $x < r_{d-1}$, then $|p_\varepsilon(x)| \geq |p(x)|$, with equality only at the common roots.

b. If $x > r_{d+1}$, then $|p_\varepsilon(x)| \leq |p(x)|$, with equality only at the common roots.

c. If $r_{d-1} < x < r_{d+1}$ and $p'(r_{d-1}) > 0$, then $p_\varepsilon(x) > p(x)$. The reverse inequality is true if $p'(r_{d-1}) < 0$.

Proof. Let $p(x) = \prod_{i=1}^{n}(x - r_i)$ be a degree n polynomial where $r_1 < r_2 < \cdots < r_n$. Let r_d be an interior root. Assume we drag r_d to the right by ε units, where $0 < \varepsilon < r_{d+1} - r_d$, such that we produce a new polynomial $p_\varepsilon(x) = (x - r_d - \varepsilon)\prod_{i \neq d}(x - r_i)$. Note that p and p_ε only intersect at values of $x = r_i$, $i \neq d$, and both take on the value of zero at these points. Further, since the degree of $p - p_\varepsilon$ is $(n-1)$, p and p_ε cannot intersect in (r_{d-1}, r_{d+1}). Also, it is important to recognize that for any two polynomials that share zeros at the endpoints of an interval and do not intersect in between the endpoints, the polynomial with the greater derivative at the left endpoint will have the greater value throughout the entire interval.

Notice that

$$p'(x) = \sum_{j=1}^{n} \prod_{i \neq j}(x - r_i),$$

and

$$p'_\varepsilon(x) = (x - r_d - \varepsilon) \sum_{j=1}^{n-1} \prod_{i \neq j,d}(x - r_i) + \prod_{i \neq d}(x - r_i).$$

Evaluating both p and p_ε at r_i, $i \neq d$, we see

$$p'(r_i) = (r_i - r_d) \prod_{j \neq i,d}(r_i - r_j),$$

while

$$p'_\varepsilon(r_i) = (r_i - r_d - \varepsilon) \prod_{j \neq i,d}(r_i - r_j).$$

Then, if $i < d$, we have $|r_i - r_d| < |r_i - r_d - \varepsilon|$, which makes $|p'(r_i) < p'_\varepsilon(r_i)|$. It then follows that $|p(x)| \leq |p_\varepsilon(x)|$ for all $x < r_{d-1}$.

Similarly, if $i > d$, $|r_i - r_d| > |r_i - r_d - \varepsilon|$, making $|p'(r_i) > p'_\varepsilon(r_i)|$ and $|p(x)| \geq |p_\varepsilon(x)|$ for all $x > r_{d+1}$. This completes parts (a.) and (b.) of the proof.

To prove part (c.) we must consider the fact that $p'(r_{d-1})$ can be negative or positive. Clearly, $(d-1) < d$, and we have already shown that $|p'(r_{d-1})| < |p'_\varepsilon(r_{d-1})|$. Therefore, if $p'(r_{d-1})$ is negative, then $p'_\varepsilon(r_{d-1})$ must be more negative, and will therefore make $|p(x)| \geq |p_\varepsilon(x)|$ on the entire interval (r_{d-1}, r_{d+1}). Similarly, if $p'(r_{d-1})$ is positive, then $p'_\varepsilon(r_{d-1})$ must be more positive, and will therefore make $|p(x)| \leq |p_\varepsilon(x)|$ over the given interval. This verifies part (c.) of the theorem, and thus completes our proof.

□

The Polynomial Root Dragging Theorem was proven in 1993, while the Polynomial Root Squeezing Theorem was proven in 2008. However, in 1967 Gideon Peyser [12] proved something with striking similarities to the Root Dragging Theorem, and he proved these things using an analysis that is very similar to the one presented above in the Root Squeezing Theorem. En route to proving that there are upper and lower bounds for a polynomial's critical numbers, he proved that critical numbers were affected by both moving and omitting exterior roots. These results were in the form of lemmas that were used in proving his theorem about the bounds of critical numbers, but they did not receive the acknowledgment that Anderson's did. They were, however, quite a profound way to analyze polynomials, and clearly influenced the proof of the Root Squeezing Theorem. Following are the proofs of the theorem as well as the lemmas that Peyser used to establish upper and lower bounds for the critical numbers of a polynomial with all real roots. In the proof, all polynomials will be written in factored form, with the roots in increasing order. That is, for a polynomial $p(x)$, we will write

$$p(x) = (x - a_1) \cdots (x - a_n)$$

and

$$p'(x) = (x - c_1) \cdots (x - c_{n-1}),$$

where a_i and c_i are real numbers and $a_i \leq c_i \leq a_{i+1}$.

Theorem 5. *If a polynomial $p(x)$ has only real roots $a_1, a_2, \ldots, a_n$, and if $a_k < a_{k+1}$, then the unique root c_k of $p'(x)$ between a_k and a_{k+1} satisfies the inequality*

$$a_k + \frac{a_{k+1} - a_k}{n - k + 1} \leq c_k \leq a_{k+1} - \frac{a_{k+1} - a_k}{k + 1}.$$

To prove this theorem, it will be helpful to prove four lemmas concerning the roots of $p(x)$ and $p'(x)$. We first consider what happens when we omit the extreme left or right roots of $p(x)$:

Lemma 1. *Let $q(x) = (x - a_1) \cdots (x - a_{n-1})$ and $q'(x) = (n-1)(x-d_1) \cdots (x-d_{n-2})$. Then $d_k \geq c_k$ for $1 \leq k \leq n-2$.*

This lemma means that when considering the polynomial $q(x)$, obtained when we remove the last root of $p(x)$, all of the roots of $q'(x)$ are greater than or equal to the corresponding roots of $p'(x)$.

Proof of Lemma 1. Let $q(x) = (x - a_1) \cdots (x - a_{n-1})$ and $q'(x) = (n - 1)(x - d_1) \cdots (x - d_{n-2})$. Then we have $p(x) = q(x)(x - a_n)$. Differentiating this yields

$$p'(x) = q'(x)(x - a_n) + q(x). \qquad (10)$$

Now consider the two consecutive roots a_k and a_{k+1}. We know that for any value of x such that $a_k < x < a_{k+1}$, $p(x)$ will be either strictly positive or strictly negative. Without loss of generality, we may assume that $p(x) > 0$ on this interval. Since $(x - a_n) < 0$, we can see that $q(x) < 0$ $a_k < x < a_{k+1}$. Further, we know that since $q(x)$ is negative, as x increases over this given interval, $q'(x)$ will change sign from negative to positive one time. We know that $p'(c_k) = 0$. Therefore, we can see from (10) that $q'(c_k)(c_k - a_n) + q(c_k) = 0$. Since $q(c_k) < 0$ and $(c_k - a_n) < 0$, we can then conclude that $q'(c_k) < 0$. Hence, since $q'(x)$ will change from negative to positive at d_k, we can conclude that $c_k < d_k$. Thus, we have completed our proof of Lemma (1).

In a similar fashion, we may prove the following:

Lemma 2. *Let $r(x) = (x - a_2) \cdots (x - a_n)$ and $r'(x) = (n - 1)(x - e_1) \cdots (x - e_{n-2})$. Then $e_k \leq c_{k+1}$ for $1 \leq k \leq n - 2$.*

This lemma means that when considering the polynomial $r(x)$, obtained when we remove the first root of $p(x)$, all of the roots of $r'(x)$ are less than or equal to the corresponding roots of $p'(x)$.

We next consider what happens when we move the leftmost or rightmost roots to the right or left, respectively:

Lemma 3. *Let $s(x) = (x - (a_1 + \varepsilon))(x - a_2) \cdots (x - a_n)$ where $\varepsilon \geq 0$ is such that $a_1 + \varepsilon \leq a_{n-1}$ and let $s'(x) = n(x - f_1) \cdots (x - f_{n-1})$. Then $f_{n-1} \geq c_{n-1}$.*

This lemma shows that if we increase the leftmost root of $p(x)$ to obtain a new polynomial $s(x)$, then the rightmost root of $s'(x)$ is greater than or equal to the rightmost root of $p'(x)$.

Proof of Lemma 3. Let $s(x) = (x-(a_1+\varepsilon))(x-a_2) \cdots (x-a_n)$ where $\varepsilon \geq 0$ is such that $a_1 + \varepsilon \leq a_{n-1}$ and let $s'(x) = n(x - f_1) \cdots (x - f_{n-1})$. We will assume that $a_{n-1} < a_n$. Then we have

$$\begin{aligned} s(x) &= [(x - a_1)(x - a_2) \cdots (x - a_n)] \\ &\quad - [\varepsilon(x - a_2)(x - a_3) \cdots (x - a_n)] \\ &= p(x) - \varepsilon r(x), \end{aligned}$$

where $r(x) = (x - a_2)(x - a_3) \cdots (x - a_n)$, as discussed in Lemma (2). Without loss of generality, we may assume that $p(x) < 0$ for $a_{n-1} < x < a_n$. Then, since $x - a_1 > 0$ on this interval, we know that $r(x) < 0$, and thus, $s(x) < 0$ for all x values within this interval. Differentiating $s(x)$, we obtain

$$s'(x) = p'(x) - \varepsilon r'(x). \qquad (11)$$

Evaluating (11) at $x = c_{n-1}$, since $p'(c_{n-1}) = 0$ we see that $s'(c_{n-1}) = -\varepsilon r'(c_{n-1})$. Since e_{n-2} is the greatest root of $r'(x)$, we may apply Lemma (2) to see that c_{n-1} is greater than every individual root of $r'(x)$. Therefore, $r'(c_{n-1})$ is a product of positive terms, and thus, is positive. Hence, $s'(c_{n-1}) < 0$. Again, as in the proof of Lemma (1), since $s'(x)$ changes sign from negative to positive values exactly one time in the interval from a_{n-1} to a_n, and $s'(f_{n-1}) = 0$, it follows that $c_{n-1} < f_{n-1}$.

In a similar fashion, we may prove the following:

Lemma 4. *Let* $t(x) = (x - a_1) \cdots (x - a_{n-1})(x - (a_n - \varepsilon))$ *where* $\varepsilon \geq 0$ *is such that* $a_n - \varepsilon \geq a_2$ *and let* $t'(x) = (n - 1)(x - g_1) \cdots (x - g_{n-2})$. *Then* $g_1 \leq c_1$.

This lemma shows that if we decrease the rightmost root of $p(x)$ to obtain a new polynomial $t(x)$, then the leftmost root of $t'(x)$ is less than or equal to the leftmost root of $p'(x)$.

We may now use these tools to prove our main theorem. The result of the theorem shows us that critical numbers cannot be found within certain intervals nearby the roots.

Proof. Let $p(x)$ be a polynomial of degree n that has only real roots $a_1, a_2, \ldots, a_n$, such that $a_k < a_{k+1}$ for all $k < n$, and critical numbers $c_1, c_2, \ldots, c_{n-1}$ such that $c_k < c_{k+1}$ for all $k < n - 1$. We will prove that the unique root c_k of $p'(x)$ between a_k and a_{k+1} satisfies the inequality

$$a_k + \frac{a_{k+1} - a_k}{n - k + 1} \leq c_k \leq a_{k+1} - \frac{a_{k+1} - a_k}{k + 1}.$$

Consider the polynomial $w(x) = (x - a_k)^k (x - a_{k+1})$. Differentiating $w(x)$, we obtain

$$\begin{aligned} w'(x) &= k(x - a_k)^{k-1}(x - a_{k+1}) + (x - a_k)^k \\ &= (x - a_k)^{k-1}[k(x - a_{k+1}) + (x - a_k)] \\ &= (x - a_k)^{k-1}(kx - ka_{k+1} + x - a_k) \\ &= (x - a_k)^{k-1}(x(k + 1) - ka_{k+1} - a_k). \end{aligned}$$

Therefore, $w'(x)$ has a root

$$c = \frac{ka_{k+1} + a_k}{k + 1} = a_{k+1} - \frac{a_{k+1} - a_k}{k + 1}.$$

We may now consider what happens when we take $p(x)$ and discard the roots $a_{k+2}, a_{k+3}, \ldots, a_n$, and drag $a_1, a_2, \ldots, a_{k-1}$ to a_k. The resulting polynomial is $w(x)$, and by applying Lemmas (1) and (3), we can conclude

$$c_k \leq c = a_{k+1} - \frac{a_{k+1} - a_k}{k + 1}.$$

Similarly, by considering the polynomial $y(x) = (x - a_k)(x - a_{k+1})^{n-k}$ and repeatedly applying Lemmas (2) and (4) to $p(x)$ to obtain $y(x)$, we can see that

$$c_k \geq a_k + \frac{a_{k+1} - a_k}{n + 1 - k}.$$

Thus, we have proven that the unique root c_k of $p'(x)$ between a_k and a_{k+1} satisfies the inequality

$$a_k + \frac{a_{k+1} - a_k}{n - k + 1} \leq c_k \leq a_{k+1} - \frac{a_{k+1} - a_k}{k + 1}.$$

□

Here we get our first glimpse of what are known as Bernstein polynomials, as given by $w(x) = (x - a_k)^k (x - a_{k+1})$ and $y(x) = (x - a_k)(x - a_{k+1})^{n-k}$. There will be much more on this very special family of polynomials later, but it is worth noting that $w(x)$ will maximize the k-th critical number for a degree $(k + 1)$ polynomial. Similarly, $y(x)$ will minimize the first critical number for a degree $(n - k + 1)$ polynomial. This makes a lot of sense when we think about it with respect to the Polynomial Root Dragging Theorem. Since all critical numbers will move in the direction of a root that is dragged, if we are trying to maximize or minimize a critical number, dragging all interior roots to the endpoints is a very intuitive solution.

In Lemmas 1-4, Peyser only addresses what happens when we omit or move exterior roots. The Polynomial Root Dragging Theorem extends the moving of roots to the interior. Here, using a very similar analysis to Peyser's, we extend this idea of omitting roots to see what happens to critical numbers when we remove interior roots. To our knowledge, this is a previously unanswered question. The proof of this lemma will follow the same assumptions presented in the previous theorem.

Lemma 5. *Let* $v(x) = (x - a_1) \cdots (x - a_{k-1})(x - a_{k+1}) \cdots (x - a_n)$ *and* $v'(x) = (n-1)(x - h_1) \cdots (x - h_{n-2})$ *where* $a_i \leq h_i \leq a_{i+1}$ *for* $i < k - 1$ *and* $a_i \leq h_{i-1} \leq a_{i+1}$ *for* $i > k$. *Then for every* $j < k - 1$, $c_j < h_j$ *and for every* $j > k$, $c_j > h_{j-1}$.

This lemma means that if $v(x)$ is obtained by omitting an interior root, r_k, of $p(x)$, then the first $k - 2$ roots of $v'(x)$ will be greater than or equal to the first $k - 2$ roots of $p'(x)$ and the last $n - k - 1$ roots of $v'(x)$ will be less than or equal to the last $n - k - 1$ roots of $p'(x)$.

Proof. Let $v(x) = (x - a_1) \cdots (x - a_{k-1})(x - a_{k+1}) \cdots (x - a_n)$ and $v'(x) = (n - 1)(x - h_1) \cdots (x - h_{n-2})$ where $a_i \leq h_i \leq a_{i+1}$ for $i < k - 1$ and $a_i \leq h_{i-1} \leq a_{i+1}$ for $i > k$. Then we have $p(x) = v(x)(x - a_k)$. Differentiating $p(x)$ gives

$$p'(x) = v'(x)(x - a_k) + v(x).$$

We know that for any value of x such that $a_j < x < a_{j+1} < a_k$, $p(x)$ will be either strictly positive or strictly negative. Without loss of generality, we may assume that $p(x) > 0$. Then, since $(x - a_k) < 0$, we can see that $v(x)$ must be negative for $a_j < x < a_{j+1}$. Further, $v'(x)$ will change sign from negative to positive exactly once over this interval. We know $p'(c_j) = 0$, so we have

$$v'(c_j)(c_j - a_k) + v(c_j) = 0.$$

Since $v(c_j) < 0$ and $(c_j - a_k) < 0$, we can see that $v'(c_j)$ must be negative. Therefore, because $v'(h_j) = 0$, we can conclude that $c_j < h_j$.

We may now consider an interval such that $a_k < a_m < x < a_{m+1}$. Again, let us assume without loss of generality that $p(x)$ is positive on this interval. Then, since $(x - a_k) > 0$, we know that $v(x)$ must be positive on the interval. Further, $v'(x)$ will change signs from positive to negative exactly once over this interval. We know that $p'(c_m) = 0$, so we have

$$v'(c_m)(c_m - a_k) + v(c_m) = 0.$$

Since $v(c_m) > 0$ and $(c_m - a_k) > 0$, we can see that $v'(c_m)$ must be negative. Therefore, since $v'(h_{m-1}) = 0$, we can conclude that $h_{m-1} < c_m$. This is exactly what we intended to show.

□

The establishment of bounds on the critical numbers of a polynomial is another very different kind of problem that the geometry of polynomials endeavors to understand. How can we estimate the value of other quantities pertaining to a polynomial's characteristics in terms of its roots and its degree? Peyser's results are a great example of this. They did, however, go relatively unnoticed for some time. In 1995, Peter Andrews [3] proved the same result of the bounds of a critical number in a different way, though he seemed unaware of Peyser's results. Then, in 2008 Aaron Melman [9] made an improvement on both of the previous theorems by a different argument through a new perspective. By considering the effect that the multiplicities of the adjacent, distinct roots have on the critical number lying between them, he made the possible interval where the critical numbers may be found slightly smaller. Note that Melman's approach to solving this problem employs a very creative application of the logarithmic derivative.

Theorem 6. *Let an nth degree polynomial $p(x)$ be given with only real roots r_i, $i = 1, 2, \ldots, n$, where $r_i \leq r_{i+1}$. If $r_k < r_{k+1}$ and the multiplicities of r_k and r_{k+1} are m_k and m_{k+1}, respectively, then the unique root c_k of $p'(x)$ between r_k and r_{k+1} satisfies the inequalities*

$$r_k + \frac{m_k}{n - k + m_k}(r_{k+1} - r_k) \leq c_k \leq r_{k+1} - \frac{m_{k+1}}{k + m_{k+1}}(r_{k+1} - r_k).$$

Proof. Let $p(x)$ be a degree n polynomial with only real roots r_i, $i = 1, 2, \ldots, n$, such that $r_i \leq r_{i+1}$. Further, let r_k and r_{k+1} be distinct roots with multiplicities m_k and m_{k+1}, respectively. We will prove that the unique root c_k of $p'(x)$ between r_k and r_{k+1} satisfies the above inequalities. We know that for some constant a,

$$p(x) = a(x - r_1)(x - r_2) \cdots (x - r_n).$$

Since $r_k < r_{k+1}$, we know that c_k is not a root of $p(x)$. In other words, $p(c_k) \neq 0$. Therefore, we can take the logarithmic derivative and evaluate it at c_k to obtain

$$\frac{p'(c_k)}{p(c_k)} = \sum_{i=1}^{n} \frac{1}{c_k - r_i} = 0.$$

We can then separate this sum into two pieces as follows:

$$\sum_{i=1}^{n} \frac{1}{c_k - r_i} = \sum_{i=1}^{k} \frac{1}{c_k - r_i} + \sum_{i=k+1}^{n} \frac{1}{c_k - r_i} = 0.$$

Subtracting $\sum_{i=k+1}^{n} \frac{1}{c_k - r_i}$ from both sides of the right-hand equation, we find that

$$\sum_{i=1}^{k} \frac{1}{c_k - r_i} = \sum_{i=k+1}^{n} \frac{1}{r_i - c_k}. \tag{12}$$

Note that for all $j \leq k$ we have $r_j < c_k$, and for all $j > k$ we have $r_j > c_k$. Therefore, all terms in equation (12) are positive. Since r_k is the largest root less than c_k, the greatest value that the left hand side of equation (12) can obtain is $\frac{k}{c_k - r_k}$. Further, since r_{k+1} has multiplicity m_{k+1}, the first m_{k+1} terms on the right side of equation (12) are identical. Since all of the terms are positive, we then know that the smallest value that the right hand side can obtain is $\frac{m_{k+1}}{r_{k+1} - c_k}$. This yields

$$\frac{k}{c_k - r_k} \geq \frac{m_{k+1}}{r_{k+1} - c_k}. \tag{13}$$

Similarly, since the multiplicity of r_k is m_k, we know that the last m_k terms of the left hand side of equation (12) are identical. So the smallest possible value that the left hand side of the equation obtains is $\frac{m_k}{c_k - r_k}$. Further, since r_{k+1} is the least root greater than c_k, and since there are $n - k$ terms on the right side of equation (12), the greatest value that the right hand side can obtain is $\frac{n-k}{r_{k+1} - c_k}$. This yields

$$\frac{m_k}{c_k - r_k} \leq \frac{n - k}{r_{k+1} - c_k}. \tag{14}$$

For notational purposes, let $\Delta = r_{k+1} - r_k$, $\Delta_k = c_k - r_k$, and $\Delta_{k+1} = r_{k+1} - c_k$. We can then rewrite inequalities (13) and (14) as

$$\frac{k}{\Delta_k} \geq \frac{m_{k+1}}{\Delta_{k+1}}$$

and

$$\frac{m_k}{\Delta_k} \leq \frac{n - k}{\Delta_{k+1}}.$$

With a little algebra we then obtain

$$\Delta_k \leq \frac{k}{m_{k+1}} \Delta_{k+1}$$

and

$$\Delta_{k+1} \leq \frac{n-k}{m_k} \Delta_k.$$

We know that $\Delta_k + \Delta_{k+1} = \Delta$. Therefore, we have

$$\Delta \leq \frac{k}{m_{k+1}} \Delta_{k+1} + \Delta_{k+1}$$

and

$$\Delta \leq \frac{n-k}{m_k} \Delta_k + \Delta_k.$$

Removing the Δ notation yields

$$r_{k+1} - r_k \leq \frac{k}{m_{k+1}} (r_{k+1} - c_k) + r_{k+1} - c_k$$

and

$$r_{k+1} - r_k \leq \frac{n-k}{m_k} (c_k - r_k) + c_k - r_k.$$

Finally, solving these inequalities for c_k, we obtain

$$c_k \leq r_{k+1} - \frac{m_{k+1}}{k + m_{k+1}} (r_{k+1} - r_k)$$

and

$$c_k \geq r_k + \frac{m_k}{n - k + m_k} (r_{k+1} - r_k).$$

Combining these last two inequalities gives us the desired result that

$$r_k + \frac{m_k}{n-k+m_k}(r_{k+1} - r_k) \leq c_k \leq r_{k+1} - \frac{m_{k+1}}{k + m_{k+1}}(r_{k+1} - r_k),$$

and thus completes our proof.

□

Therefore, just by considering the multiplicities of roots a factor, Melman has successfully narrowed the possible interval in which a critical number may live. In his paper, he provides an example that is helpful in illustrating the value of this. For a polynomial $q(x) = (x-1)(x-2)^4(x-3)^2(x-4)$, according to Peyser and Andrew's method, the fifth critical number would fall between $2\frac{1}{4}$ and $2\frac{5}{6}$. However, by considering the multiplicities, Melman notes that the same root must actually fall between $2\frac{4}{7}$ and $2\frac{5}{7}$. This is quite an improvement over the previous theorems.

The following is one final example of the kind of problems we are interested in. It is a very short proof by Răzvan Gelca of a theorem dealing with the separation of zeros of a polynomial. Again, we see the use of the logarithmic derivative. In order to prove the theorem, we must make a few assumptions. For a polynomial $f(x) = \prod_{i=1}^{n}(x - x_i)$, with distinct real roots $x_1 < x_2 < \cdots < x_n$, we let $d = \delta(f) = \min_i(x_{i+1} - x_i)$ and $g(x) = f'(x)/f(x) = \sum_{i=1}^{n} 1/(x - x_i)$. If k is a real number, then the roots of the polynomial $f' - kf$ are also real and distinct.

Theorem 7. *If some j, y_0 and y_1 satisfy $y_0 < x_j < y_1 \leq y_0 + d$, then y_0 and y_1 are not zeros of f and $g(y_0) < g(y_1)$.*

Proof. Let j, y_0 and y_1 satisfy $y_0 < x_j < y_1 \leq y_0 + d$. We will prove that y_0 and y_1 are not zeros of f and $g(y_0) < g(y_1)$. It follows directly from the hypothesis that for all i, $y_1 - y_0 \leq d \leq x_{i+1} - x_i$. Then for $1 \leq i \leq j-1$, we have

$$y_1 - y_0 \leq x_{i+1} - x_i.$$

Multiplying by (-1) yields

$$y_0 - y_1 \geq x_i - x_{i+1}.$$

A little arithmetic gives us

$$y_0 - x_i \geq y_1 - x_{i+1}.$$

Since $x_{i+1} \leq x_j$, we can then conclude that

$$y_0 - x_i \geq y_1 - x_{i+1} > 0.$$

Therefore, we have

$$\frac{1}{(y_0 - x_i)} \leq \frac{1}{(y_1 - x_{i+1})}. \tag{15}$$

Similarly, for $j \leq i \leq n-1$ we have $y_1 - x_{i+1} \leq y_0 - x_i < 0$, and we still have the same inequality achieved in (15). Obviously, the inequality

$$y_0 - x_n < 0 < y_1 - x_1 \tag{16}$$

is true, so we have

$$\frac{1}{(y_0 - x_n)} < 0 < \frac{1}{y_1 - x_1}. \tag{17}$$

We have now shown that $y_0 - x_i \neq 0$ and $y_1 - x_i \neq 0$ for all $i \in \{1, 2, \ldots, n\}$. Therefore, neither y_0 nor y_1 are zeros of f. Further, adding (16) and (17) shows that $g(y_0) < g(y_1)$.

□

We have now seen a broad range of problems concerning the geometry of polynomials. These fundamental theorems serve a few purposes. They provide us with tools that we may use in further investigating related problems, as well as an illustration of the kind of problems that the geometry of polynomials is concerned with answering. We now have a strong enough foundation to begin exploring some problems dealing with extremality.

3 Special Polynomials

There are two different families of polynomials that will be especially important to be familiar with in order to discuss extremal problems. These famous families are known as Chebyshev polynomials and Bernstein polynomials.

Chebyshev polynomials come in four different kinds. There are two that we are interested in discussing: Chebyshev polynomials of the first kind, and Chebyshev polynomials of the second kind. A degree n Chebyshev polynomial of the first kind is defined recursively as follows:

$$T_0(x) = 1 \qquad T_1(x) = x$$
$$T_{n+1}(x) = 2x \cdot T_n(x) - T_{n-1}(x) \qquad (n \geq 1).$$

These polynomials are equioscillatory, meaning the deviation of the curve from the x-axis at the critical values, and at the end points of the interval, is equal. These polynomials also have a closed form expression, which is much more useful in application. The following is a short proof highlighting this, and it is found in a Numerical Analysis text by Ward Cheney and David Kincaid [6].

Theorem 8. *For x in the interval $[-1, 1]$, the Chebyshev polynomials of the first kind have this closed-form expression*

$$T_n(x) = \cos(n \cos^{-1} x), \quad (n \geq 0).$$

Proof. To prove this, we must first recall the addition formula for the cosine:

$$\cos(A + B) = \cos A \cos B - \sin A \sin B.$$

From this, we obtain

$$\begin{aligned} \cos((n+1)\theta) &= \cos(\theta n + \theta) \\ &= \cos\theta \cos n\theta - \sin\theta \sin n\theta \end{aligned} \tag{18}$$

and

$$\begin{aligned} \cos((n-1)\theta) &= \cos(\theta n - \theta) \\ &= \cos\theta \cos n\theta + \sin\theta \sin n\theta. \end{aligned} \tag{19}$$

We may now add (18) and (19) together, and we get

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos\theta \cos n\theta$$

which yields

$$\cos((n+1)\theta) = 2\cos\theta\cos n\theta - \cos((n-1)\theta). \tag{20}$$

Now, if we let $\theta = \cos^{-1} x$ and $x = \cos\theta$, we can see that equation (20) shows that the functions f_n defined by

$$f_n(x) = \cos(n \cos^{-1} x)$$

follow the system of equations:

$$f_0(x) = 1$$
$$f_1(x) = x$$
$$f_{n+1}(x) = 2x f_n(x) - f_{n-1}(x), \qquad (n \geq 1).$$

Therefore, we have proven that $f_n(x) = T_n(x)$ for all n.

□

Using this closed form expression makes it easy to see that Chebyshev polynomials of the first kind are indeed equioscillatory on the interval $[-1, 1]$.

For Chebyshev polynomials of the second kind, we will start by introducing the closed form expression. A degree n Chebyshev polynomial of the second kind is defined on the interval $[-1, 1]$ as

$$U_n(x) = \frac{\sin((n+1)\arccos(x))}{\sqrt{1-x^2}}.$$

These polynomials do follow the same recursive definition as the Chebyshev polynomials of the first kind. That is,

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \qquad (n \geq 1).$$

Further, the first two kinds of Chebyshev polynomials have the very nice relationship that

$$U_n(x) = \frac{1}{n} T'_{n+1}(x).$$

These polynomials arise as minimizers in a couple of our extremal problems, as we will soon explore further. Figure 1 shows an example of each of the two kinds of Chebyshev polynomials that we are interested in.

On the flip side of extremal problems, Bernstein polynomials, as briefly introduced in Section 2, frequently arise as maximizers. A degree n Bernstein polynomial is defined as

$$B_{i,n}(x) = \binom{n}{i} x^i (1-x)^{n-i}.$$

This is a family of polynomials. To be a member of this family, a polynomial must have all of its roots at 0 or 1. However, since we are interested in monic polynomials with all real zeros in the interval $[-1, 1]$, we have scaled the Bernstein polynomials to meet our requirements. That is, for our purposes, we define Bernstein polynomials in the following way:

$$B_{i,n}(x) = (x+1)^i (x-1)^{n-i}.$$

Notice that neither of the forms provided here match the forms of the functions $w(x)$ and $y(x)$ from Theorem 5 in Section 2. This just shows that Bernstein Polynomials can be defined on any interval, and the only real requirement is that we distribute all of the roots between the endpoints of the interval. So we are interested in the monic Bernstein polynomials formed by placing all of the roots at either 1 or -1. Note that for any given n, there are $n+1$ degree n Bernstein polynomials. Understanding why these polynomials so frequently maximize different properties is of particular interest

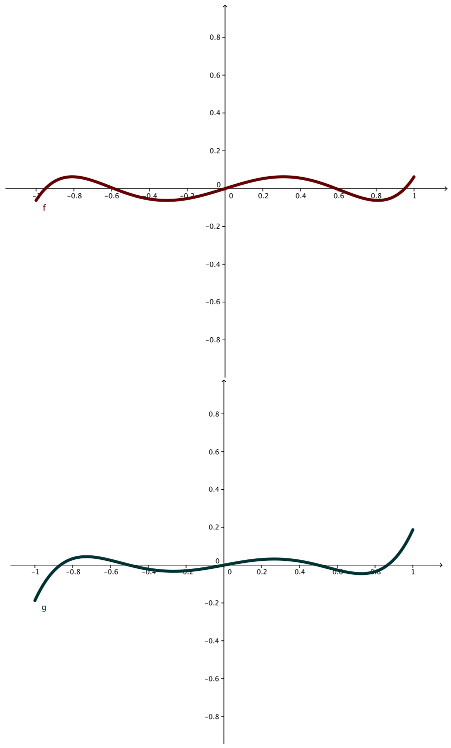

Figure 1: Degree 5 monic Chebyshev polynomials of the first and second kind

to us, and this will be the focus of much of the rest of the paper.

Figure 2 provides an example of a monic degree 5 Bernstein polynomial with four roots at 1 and one root at -1.

An in depth look at many problems concerning these two families of polynomials is to follow. Ultimately, we endeavor to further understand the phenomena of these families and their maximal and minimal qualities.

4 Extremal Problems

We now begin our discussion of problems concerning polynomials that possess maximal and minimal properties. There are many important results and concepts that must first be introduced. One important concept when dealing with these extremal problems is the idea of norms. Norms are types of functions that measure some aspect of another function. A norm is analogous to the absolute value function for real numbers, as it will return a non-negative real number that may be thought of as "size." There are other functions that we may use to measure properties as well, and these different ways to measure properties are the only way we can make sense of discussing maximal and minimal polynomials.

Specifically, we will be focusing on the supremum norm and the L^1 norm. The supremum norm of a function $p(x)$ is defined by

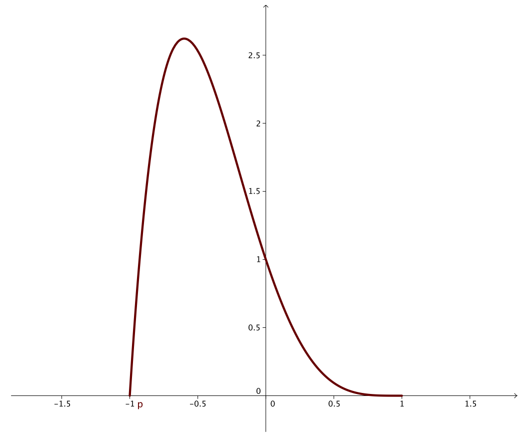

Figure 2: Degree 5 monic Bernstein Polynomial

$$||p(x)||_\infty = \max_{x\in[-1,1]} |p(x)|.$$

The L^1 norm of a function $p(x)$ is defined as

$$||p(x)||_1 = \int_{-1}^{1} |p(x)|dx.$$

Using these norms, and other functions, we will present a few examples of extremal problems in which the answer is either Chebyshev polynomials or Bernstein polynomials. Chebyshev polynomials have been shown to be the minimizer of the supremum norm and the L^1 norm, as we will see. We will also show that Bernstein polynomials are the maximizers of the supremum norm. A remaining question regards the maximizer of the L^1 norm. In this section we will also explore four other properties that are maximized by Bernstein polynomials.

To begin, we will provide a proof that the supremum norm is minimized by Chebyshev polynomials of the first kind. In light of Theorem 4, the fact that Chebyshev polynomials are equioscillatory makes them a good candidate for the supremum norm minimizer: since dragging roots makes the deviation grow in some places and shrink in others, it makes sense intuitively that the minimum maximum deviation must occur for a polynomial where the deviation at the critical points is equal. As it turns out, the leading coefficient for the degree n Chebyshev polynomial of the first kind is 2^{n-1}. Therefore, the monic Chebyshev polynomial of the first kind is $2^{1-n}T_n$. The following proof was presented in a Numerical Analysis text by Ward Cheney and David Kincaid [6].

Theorem 9. *If p is a monic polynomial of degree n, then*

$$||p||_\infty = \max_{-1\leq x\leq 1} |p(x)| \geq 2^{1-n}.$$

Proof. Let p be a monic polynomial of degree n. We will prove that $||p||_\infty = \max_{-1\le x\le 1}|p(x)| \ge 2^{1-n}$. To do this, we will argue by contradiction. That is, we will assume that for $-1 \le x \le 1$,

$$|p(x)| < 2^{1-n}.$$

Let $q = 2^{1-n}T_n$ and $x_i = \cos(i\pi/n)$, where $0 \le i \le n$. We may now use properties of Chebyshev polynomials to see that q is a monic polynomial of degree n and

$$(-1)^i p(x_i) \le |p(x_i)| < 2^{1-n} = (-1)^i q(x_i).$$

Therefore, we may use algebra to see that for $0 \le i \le n$,

$$(-1)^i[q(x_i) - p(x_i)] > 0.$$

From this we see that the polynomial $q - p$ oscillates in sign $(n + 1)$ times between -1 and 1. However, as the leading terms in $q-p$ will cancel out, $q-p$ cannot have degree higher than $n-1$. Therefore, $q-p$ could not possibly oscillate $(n+1)$ times between -1 and 1. Hence, we have reached our contradiction, and may conclude that if p is a monic polynomial, then $||p||_\infty = \max_{-1\le x\le 1}|p(x)| \ge 2^{1-n}$, as we desired.

□

Now, we present the theorem that proved the L^1 norm is minimized by Chebyshev Polynomials of the Second Kind. This proof is presented in the book *Topics in Polynomials: Extremal Problems, Inequalities, and Zeros*, by G.V. Milovanovic [10]. There was a lemma necessary to prove this, and here we present both the lemma and the theorem. We have omitted the proofs of these, as for our purposes, the results are only being used as an illustration of the minimality of Chebyshev polynomials.

Lemma 6. *Let*

$$U_n(x) = \frac{\sin((n+1)\arccos(x))}{\sqrt{1-x^2}}$$

be the n-th Chebyshev polynomial of the second kind. Then

$$I_{n,k} = \int_{-1}^{1} x^k \operatorname{sgn} U_n(x)dx = \begin{cases} 0, & \text{if } 0 \le k \le n-1 \\ 2^{1-n}, & \text{if } k = n. \end{cases} \tag{21}$$

This lemma is then used to prove the following theorem:

Theorem 10. *Let $p(x) = \sum_{i=0}^{n} a_i x^i$, with $a_n = 1$, be an arbitrary monic polynomial of degree n. Then*

$$||p(x)||_1 \ge ||\hat{U}_n||_1 = 2^{1-n},$$

with equality only if $p(x) = \hat{U}_n(x)$, where $\hat{U}_n$ is the monic Chebyshev polynomial of the second kind of degree n. In other words, $\hat{U}_n(x) = 2^{-n}U_n(x)$.

We have now illustrated how Chebyshev polynomials arise as minimizers of certain properties. Let us now turn our focus to maximizers. By employing Theorem 4, we see that in order to maximize the deviation of the curve from the x-axis near a specific critical value, we must drag all of the roots as far away from that point as possible. In our case, since we are interested strictly in polynomials with roots that live in $[-1, 1]$, these maximal polynomials that we construct will turn out to be members of the Bernstein family. According to the definition of the supremum norm, we see that one of these Bernstein polynomials must maximize the supremum norm of a monic polynomial with all real zeros in the interval $[-1, 1]$. The next question, then, is which one? This question is also answered in the same paper that presented Theorem 4 [5].

Theorem 11. *Given a family of degree n Bernstein polynomials, $p_i(x) = (x+1)^i(x-1)^{n-i}$, the supremum norm of p_i will be maximized when $i = 1$ or when $i = n-1$.*

Proof. Let $p_i(x) = (x+1)^i(x-1)^{n-i}$ be the family of degree n Bernstein polynomials in the interval $[-1, 1]$. Since each p_i must attain its supremum norm at a critical point, and each p_i will only have one critical point in $(-1, 1)$, we need only look at the value of p_i at this critical point, which we will call c_i. To find c_i, we must evaluate $p_i'(x) = 0$. Note that

$$\begin{aligned} p_i'(x) \quad &= i(x+1)^{i-1}(x-1)^{n-i} \\ &\quad + (n-i)(x-1)^{n-i-1}(x+1)^i \\ &= (x-1)^{n-i-1}(x+1)^{i-1} \\ &\quad \times (i(x-1) + (n-i)(x+1)) \\ &= (x-1)^{n-i-1}(x+1)^{i-1}(nx + n - 2i). \end{aligned}$$

Therefore, for $p_i'(x) = 0$, x must either be 1, -1, or $\frac{2i-n}{n}$. Hence, $c_i = \frac{2i-n}{n}$, and the supremum norm of p_i is simply $|p_i(c_i)|$, which is

$$\begin{aligned} |p_i(c_i)| &= \left| \left(\frac{2i-n}{n} + 1\right)^i \left(\frac{2i-n}{n} - 1\right)^{n-i} \right| \\ &= \left(\frac{2^i i^i}{n^i}\right) \left| \left(\frac{2^{n-i}(i-n)^{n-i}}{n^{n-i}}\right) \right| \\ &= \left(\frac{2^n}{n^n}\right) i^i (n-i)^{n-i}. \end{aligned}$$

Now, to maximize the supremum norm for all p_i, we must find the value of i that makes the value of $|p_i(c_i)|$ the greatest. Since n will remain fixed, we can simply maximize $i^i(n-i)^{n-i}$. Instead, we may maximize the continuous function $g(t) = t^t(n-t)^{n-t}$ on $t \in [1, n-1]$. Since $g(t)$ and $\ln(g(t))$ have the same critical numbers, we can equivalently find the critical values of $f(t) = \ln(t^t(n-t)^{n-t})$. Using the chain rule, we obtain

$$\begin{aligned} f'(t) &= \ln(t) + 1 - \ln(n-t) - 1 \\ &= \ln(t/(n-t)). \end{aligned}$$

Therefore, $f'(t)$ will only be zero when $\frac{t}{n-t} = 1$. Hence, the only critical number of f, and consequently that of g, will be $t = \frac{n}{2}$. However, this value produces a minimum of g on $[1, n-1]$, and thus, g must be maximized at the endpoints. So the maximum of g and also that of $|p_i(c_i)|$ will occur when $t = 1$ or $t = n-1$, and these values will be the same because of the symmetry of $|p_i(c_i)|$.

$\square$

This paper also establishes the maximum value of the supremum norm for a degree n monic polynomial with all real zeros. Since the maximum supremum norm of p_i will be attained when $i = 1$, we can see that

$$|p_1(c_1)| = \frac{2^n}{n^n}(n-1)^{n-1} = \frac{2^n}{n}\left(\frac{n-1}{n}\right)^{n-1}.$$

Hence, for any monic degree n polynomial, p, with all real zeros in $[-1, 1]$,

$$||p||_\infty \leq \frac{2^n}{n}\left(\frac{n-1}{n}\right)^{n-1}.$$

Next, we present a conjecture that we have not yet been able to prove. It concerns the maximal L^1 norm. We are convinced that the following conjecture is true, and there is strong evidence to support this. Although we have tried many approaches to solve this, a formal proof eludes us.

Conjecture 1. *The L^1 norm of a degree n monic polynomial with all real zeros is maximized by a polynomial having the form $p(x) = (x+1)(x-1)^{n-1}$ or $p(x) = (x+1)^{n-1}(x-1)$.*

Now we offer an alternative proof to a maximal problem proven by Blagovest Sendov [14] in 2001. This is a problem related to the Sendov Conjecture, which is a notable unsolved problem.

Conjecture 2. (**Sendov Conjecture**) *If all the zeros of the polynomial $p(z) = \prod_{k=1}^{n}(z - z_k)$, $(n \geq 2)$ lie in the unit disk $D(0,1) = \{z : |z| \leq 1\}$, then for every z_k, the disk $D(z_k, 1)$ contains at least one zero of $p'(z)$.*

Before we state the theorem that we are providing an alternate proof for, we must first introduce some new notation. The problem we are answering concerns the *Hausdorff deviation* of the set of roots of a polynomial from the *convex hull* of the polynomial's critical numbers. We now present the definition of the Hausdorff deviation of a set B from a set A, which is noted $\rho(B, A)$.

Definition 4. *Given two sets A and B, the Hausdorff deviation of B from A, $\rho(B, A)$, is the supremum of the set $\{\rho(b, A) : b \in B\}$, where $\rho(b, A) = \inf\{|b - a| : a \in A\}$.*

Intuitively, the Hausdorff deviation of B from A can be thought of as finding the point in B that is furthest from the nearest point in A, and measuring the distance between the two points. It is now necessary to define the convex hull of a set A, which is denoted $H(A)$.

Definition 5. *Given a set A, the smallest closed and convex point set that contains A is called the convex hull of A, denoted $H(A)$.*

Finally, we introduce the notation $A(p)$, which represents the set of all distinct zeros of a polynomial. We are now ready to present the theorem.

Theorem 12. *For every polynomial p of degree n with all real zeros, the inequality*

$$\rho(A(p), H(p')) \leq \frac{2}{n}$$

holds.

In order to understand fully what this theorem is asking, a picture may be helpful. Figure 3 illustrates what a possible degree 5 case of this theorem would look like.

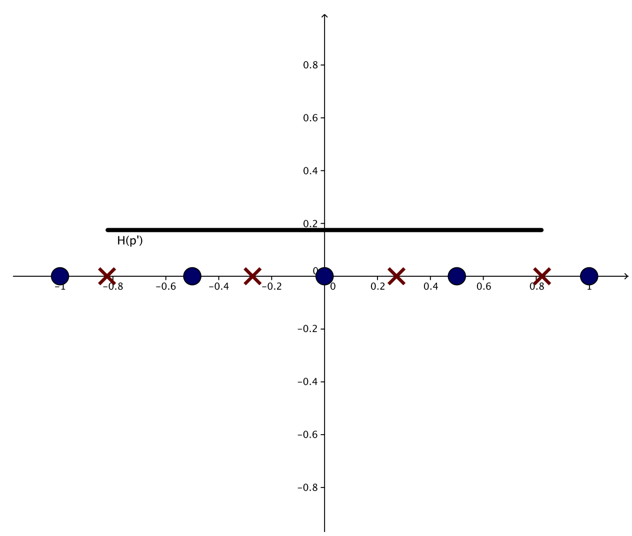

Figure 3: Degree 5 illustration of Theorem 12

In figure 3, the dots represent the roots, the "X's" represent the critical numbers, and the line above the x-axis represents the convex hull of the critical numbers. Note that in reality, the convex hull will be on the x-axis, but for sake of clarity it is shown above. The theorem is then saying that the maximum distance from the set of roots to the convex hull is $2/n$. It is important to note that all of the interior roots will also be elements of the convex hull of the critical numbers. Therefore, we are really trying to show that the distance from one of the exterior roots to the nearest critical number will never exceed $2/n$. We will prove this using a root dragging argument.

Proof. Let $p(x)$ be a degree n monic polynomial with all real zeros in the interval $[-1,1]$. We will prove that the maximum Hausdorff deviation of $A(p)$ from $H(p')$ is $2/n$. The Polynomial Root Dragging Theorem says that if we drag the interior roots of p to the right, the critical numbers of p will follow. Similarly, if we drag the interior roots to the left, the critical numbers will follow to the left. Therefore, by dragging the rightmost interior root to 1, the convex hull of the critical numbers has also moved to the right, and increased the Hausdorff deviation of $A(p)$ from $H(p')$. In order to maximize the Hausdorff deviation of $A(p)$ from $H(p')$, we may continue in this fashion and drag all of the interior roots to 1, creating a Bernstein polynomial with one root at -1 and $n-1$ roots at 1. Note that, due to the symmetry of Bernstein polynomials, $\rho(A(p), H(p'))$ would be the same if we dragged all of the interior roots to -1. So in order to maximize $\rho(A(p), H(p'))$, p must have the form of

$$p(x) = (x+1)(x-1)^{n-1},$$

or

$$p(x) = (x+1)^{n-1}(x-1).$$

So $\rho(A(p), H(p'))$ will be the distance from the critical point in $(-1,1)$ and the single non-repeated root. Again, because of the symmetry of Bernstein polynomials, this value will be the same in either case. Let us then find $\rho(A(p), H(p'))$ for $p(x) = (x+1)(x-1)^{n-1}$. We must first find the critical number of p in $(-1,1)$, which we will call c. Recall from Theorem 11 that $c = \frac{2-n}{n}$. So we have

$$\begin{aligned}\rho(A(p), H(p')) &= \left(\frac{2-n}{n}\right) - (-1)\\ &= \left(\frac{2}{n} - 1\right) + 1\\ &= \frac{2}{n}.\end{aligned}$$

Hence, the Hausdorff deviation of a degree n polynomial's roots from the convex hull of its critical numbers must be less than or equal to $2/n$, with equality only for these special Bernstein polynomials.

□

A similar extremal problem, considered by Piotr Pawlowski [11] in 1998, shows how we may maximize the distance from the centroid, A_p of a polynomial to its nearest critical number. That is, given a degree n polynomial p, with critical numbers c_k, the theorem shows how we can maximize the function

$$J(p) = \min_{1 \le k \le n-1} |A_p - c_k|.$$

Pawlowski states that $J(p)$ cannot exceed $2/3$, and that it may only reach $2/3$ for a very specific kind of Bernstein polynomial. We do not present Pawlowski's proof of this result here, for it requires some very complex mathematics. However, we do state the theorem as proved by Pawlowski, which highlights the kind of Bernstein polynomial necessary for maximizing $J(p)$. Also, we provide a more intuitive reasoning for why $J(p)$ can never exceed $2/3$ and why n must be a multiple of 3 to maximize this property.

Theorem 13. *If $p(x)$ is a polynomial with all real zeros, then*

$$J(p) \le \frac{2}{3}.$$

Equality is attained if and only if n is a multiplicity of 3 and $p(x) = (x-1)^{\frac{2n}{3}}(x+1)^{\frac{n}{3}}$ or $p(x) = (x-1)^{\frac{n}{3}}(x+1)^{\frac{2n}{3}}$.

Recall that the centroid of a polynomial is the average of the roots. Therefore, in the case of a Bernstein polynomial, $B_{i,n}(x) = (x+1)^i(x-1)^{n-i}$, the centroid is

$$\begin{aligned}A_B &= \frac{(-1)i + (n-i)}{n}\\ &= \frac{n-2i}{n}.\end{aligned}$$

Further, by using the standard method for finding critical numbers, we may see that the interior critical number of $B_{i,n}$ is $(2i-n)/n$. So the centroid of a Bernstein polynomial is exactly the opposite of the single interior critical number of the polynomial. So the distance from the centroid to the nearest endpoint will be $1 - |(n-2i)/n|$, and the distance from the centroid to the interior critical number will be $2|(n-2i)/n|$. If we do anything to increase either of these distances, the other one must decrease. Therefore, since $J(B)$ measures the smallest of these two distances, maximizing $J(B)$ will require these values to be equal. We then must have

$$1 - \left|\frac{n-2i}{n}\right| = 2\left|\frac{n-2i}{n}\right|,$$

or

$$\left|\frac{n-2i}{n}\right| = \frac{1}{3}.$$

Hence, in order to maximize $J(B)$, we must have a centroid that falls exactly at $1/3$ or $-1/3$, making $J(B) = 2/3$. Now, we can clearly see that $|(n-2i)/n|$ can only be $1/3$ when n is a multiple of 3, which explains why this is necessary for maximization of this property.

Finally, we introduce our last example of a maximal problem. This problem, as solved by Raphael Robinson [13] in 1964, deals with maximizing the span of the k-th derivative of a degree n polynomial with all real roots. The proof of this result uses the very important concept of convexity, which we will explore more thoroughly in the next section. This approach to solving this problem is very creative, but again, we will simply highlight the result here.

Theorem 14. *The span of the kth derivative of a polynomial with all real zeros, $p^{(k)}(x)$, can be maximized only when all of the roots of $p(x)$ are at the end points, $x = \pm 1$.*

Note that this problem is trivial for values of k that are greater than $n-2$, for we want $f^{(k)}$ to have more than one root. The problem is also trivial if $2k+2 \leq n$, because we could then simply place $k+1$ roots at both endpoints, and the span of the k-th derivative would be 2. Therefore, we may narrow our focus to the nontrivial cases, with $k+2 \leq n \leq 2k+1$. The proof shows that if we have a polynomial with all but one of its roots determined, the only way the span of the k-th derivative of the polynomial can be maximized is by placing the undetermined root at either 1 or -1. Then, by applying this argument to each root of the polynomial individually, it shows that the span of $p^{(k)}(x)$ can be maximized only when all of the roots of $p(x)$ are at 1 or -1. In other words, this property is only maximized by a Bernstein polynomial. However, while the evidence for an evenly distributed Bernstein polynomial is strong, it has not yet been proven which Bernstein polynomial will maximize the property.

So there is this obvious pattern in these maximal problems. As illustrated with the problems above, the family of Bernstein polynomials frequently provides us with the maximizer of a characteristic. This is the phenomenon to which we spoke in the opening of this paper, and from here, we may begin our discussion of our attempt of a general explanation.

5 Convexity

The theory of convexity tells us that any convex function over a compact, convex set must attain its maximum at an extreme point. This is precisely the idea that we believe will help us explain the Bernstein phenomenon. In order to understand this fully, there are a couple of key concepts that must be introduced.

Definition 6. *A convex set is a collection of points such that every line segment formed by connecting points in the set is contained within the set. In other words, a set A is convex if and only if given any two elements of A, a and b, and any real number α such that $0 \leq \alpha \leq 1$,*

$$\alpha \cdot a + (1-\alpha) \cdot b \in X.$$

Once we have a convex set, we may think about what are known as convex functions.

Definition 7. *A convex function is a real valued function over a convex set such that the region above the graph of the function is also a convex set. In other words, a function f is convex if and only if given any two elements of the domain, x and y, and a real number α such that $0 \leq \alpha \leq 1$,*

$$f(\alpha \cdot x + (1-\alpha) \cdot y) \leq \alpha \cdot f(x) + (1-\alpha) \cdot f(y).$$

One of the first things we must do in order to use this as a tool is develop a method in which we may think of monic polynomials with all real zeros as a vector space. To have a vector space, we need closure under multiplication and addition. The set of monic polynomials with all zeros, however, is not closed under normal function addition. It is not hard to construct a counterexample that shows that we may add two monic polynomials with all zeros and obtain a polynomial with complex zeros. For example, consider the polynomials $p(x) = x^2 - 1$ and $q(x) = x^2 + 5x + 6$. The polynomial p has roots at -1 and 1, while q has roots at -2 and -3. However, when adding p and q, we obtain a new polynomial $r = 2x^2 + 5x + 5$, which has two imaginary roots. Therefore, we must develop a new way to add and multiply polynomials. Here we introduce two new operations, which we will denote $\oplus$ and $\otimes$, and we will prove that they make the set of monic polynomials with all real zeros a vector space.

Definition 8. *Given two degree n monic polynomials with all real zeros, $p(x) = \prod_{i=1}^{n}(x - r_i)$ and $q(x) = \prod_{i=1}^{n}(x - s_i)$, we define the $\oplus$ operator by the rule*

$$p(x) \oplus q(x) = \prod_{i=1}^{n}(x - (r_i + s_i)).$$

Definition 9. *Given any degree n monic polynomials with all real zeros, $p(x) = \prod_{i=1}^{n}(x - r_i)$ and any real number, α, we define the $\otimes$ operator by the rule*

$$\alpha \otimes p(x) = \prod_{i=1}^{n}(x - \alpha r_i).$$

We can think about these somewhat unusual operations in a very natural way with respect to root-dragging. When we "add" a polynomial q to a polynomial p, we may think of q as the polynomial that tells us how much we will drag the roots of p by. Similarly, when we "multiply" p by a scalar, α, if $0 < \alpha < 1$, it is as if we are simultaneously dragging all of the roots of p towards the origin by a certain multiple of themselves; if α is greater than 1, it is as if we are dragging the roots away from the origin. When defining polynomial addition and multiplication in this way, we have the zero element x^n, which we will denote 0_P. We will now show that under these two operations, the set of monic polynomials with all real zeros is a vector space.

Theorem 15. *Under the operations $\oplus$ and $\otimes$ as defined above, the set of monic polynomials with all real zeros forms a vector space.*

Proof. Let the operations $\oplus$ and $\otimes$ be given as in definitions 8 and 9, and let 0_P be the zero element of the set of monic polynomials with all real zeros. We will prove that under these assumptions, the set of monic polynomials with all real zeros is a vector space. To do this, we must show that the following properties hold true for any degree n monic polynomials with all real zeros, p, q, and r, and for any real numbers a and b:

i. $p \oplus q = q \oplus p$

ii. $(p \oplus q) \oplus r = p \oplus (q \oplus r)$

iii $0_P \oplus p = p = p \oplus 0_P$

iv. $(-p) \oplus p = 0_P = p \oplus (-p)$

v. $0_P \otimes p = 0_P$

vi. $1 \otimes p = p$

vii. $(ab) \otimes p = a \otimes (b \otimes p)$

viii. $a \otimes (p \oplus q) = (a \otimes p) \oplus (a \otimes q)$

ix. $(a + b) \otimes p = (a \otimes p) \oplus (b \otimes p)$

Let p, q, and r be monic polynomials with all real zeros, and let a and b be real numbers. To prove these properties hold true, we will let the roots of p, q, and r be represented as p_i, q_i, and r_i, respectively. Then we have $p(x) = \prod_{i=1}^{n}(x - p_i)$, $q(x) = \prod_{i=1}^{n}(x - q_i)$, and $r(x) = \prod_{i=1}^{n}(x - r_i)$. We will prove that each of these holds true one at a time:

(i.):

$$\begin{aligned} p \oplus q &= \prod_{i=1}^{n}(x - p_i) \oplus \prod_{i=1}^{n}(x - q_i) \\ &= \prod_{i=1}^{n}(x - (p_i + q_i)) \\ &= \prod_{i=1}^{n}(x - (q_i + p_i)) \\ &= \prod_{i=1}^{n}(x - q_i) \oplus \prod_{i=1}^{n}(x - p_i) \\ &= q \oplus p, \end{aligned}$$

(ii.):

$$\begin{aligned} (p \oplus q) \oplus r &= \left(\prod_{i=1}^{n}(x - p_i) \oplus \prod_{i=1}^{n}(x - q_i)\right) \oplus \prod_{i=1}^{n}(x - r_i) \\ &= \prod_{i=1}^{n}(x - (p_i + q_i)) \oplus \prod_{i=1}^{n}(x - r_i) \\ &= \prod_{i=1}^{n}(x - (p_i + q_i + r_i)) \\ &= \prod_{i=1}^{n}(x - p_i) \oplus \prod_{i=1}^{n}(x - (q_i + r_i)) \\ &= p \oplus (q \oplus r), \end{aligned}$$

(iii.):

$$\begin{aligned} 0_P \oplus p &= \prod_{i=1}^{n}(x - 0) \oplus \prod_{i=1}^{n}(x - p_i) \\ &= \prod_{i=1}^{n}(x - (0 + p_i)) \\ &= \prod_{i=1}^{n}(x - p_i) = p \\ &= \prod_{i=1}^{n}(x - (p_i + 0)) \\ &= \prod_{i=1}^{n}(x - p_i) \oplus \prod_{i=1}^{n}(x - 0) \\ &= p \oplus 0_P, \end{aligned}$$

(iv.):

$$\begin{aligned} (-p) \oplus p &= \prod_{i=1}^{n}(x - (-p)) \oplus \prod_{i=1}^{n}(x - p_i) \\ &= \prod_{i=1}^{n}(x - ((-p) + p_i)) \\ &= \prod_{i=1}^{n}(x - 0) = 0_P \\ &= \prod_{i=1}^{n}(x - (p_i + (-p_i))) \\ &= \prod_{i=1}^{n}(x - p_i) \oplus \prod_{i=1}^{n}(x - (-p_i)) \\ &= p \oplus (-p), \end{aligned}$$

(v.):

$$\begin{aligned} 0_p \otimes p &= \prod_{i=1}^{n}(x - 0 \cdot p_i) \\ &= \prod_{i=1}^{n}(x - 0) \\ &= 0_p, \end{aligned}$$

(vi.):

$$\begin{aligned} 1 \otimes p &= 1 \otimes \prod_{i=1}^{n}(x - p_i) \\ &= \prod_{i=1}^{n}(x - 1 \cdot p_i) \\ &= \prod_{i=1}^{n}(x - p_i) \\ &= p, \end{aligned}$$

(vii.):

$$\begin{aligned}(ab)\otimes p &= (ab)\otimes\prod_{i=1}^{n}(x-p_i)\\ &= \prod_{i=1}^{n}(x-(ab)\cdot p_i)\\ &= \prod_{i=1}^{n}(x-a(b\cdot p_i))\\ &= a\otimes\prod_{i=1}^{n}(x-b\cdot p_i)\\ &= a\otimes\left(b\otimes\prod_{i=1}^{n}(x-p_i)\right)\\ &= a\otimes(b\otimes p),\end{aligned}$$

(viii.):

$$\begin{aligned}a\otimes(p\oplus q) &= a\otimes\left(\prod_{i=1}^{n}(x-p_i)\oplus\prod_{i=1}^{n}(x-q_i)\right)\\ &= a\otimes\prod_{i=1}^{n}(x-(p_i+q_i))\\ &= \prod_{i=1}^{n}(x-a(p_i+q_i))\\ &= \prod_{i=1}^{n}(x-(ap_i+aq_i))\\ &= \prod_{i=1}^{n}(x-ap_i)\oplus\prod_{i=1}^{n}(x-aq_i)\\ &= \left(a\otimes\prod_{i=1}^{n}(x-p_i)\right)\oplus\left(a\otimes\prod_{i=1}^{n}(x-q_i)\right)\\ &= (a\otimes p)\oplus(a\otimes q),\end{aligned}$$

(ix.):

$$\begin{aligned}(a+b)\otimes p &= (a+b)\otimes\prod_{i=1}^{n}(x-p_i)\\ &= \prod_{i=1}^{n}(x-(a+b)\cdot p_i)\\ &= \prod_{i=1}^{n}(x-(a\cdot p_i+b\cdot p_i))\\ &= \prod_{i=1}^{n}(x-a\cdot p_i)\oplus\prod_{i=1}^{n}(x-b\cdot p_i)\\ &= \left(a\otimes\prod_{i=1}^{n}(x-p_i)\oplus b\otimes\prod_{i=1}^{n}(x-p_i)\right)\\ &= (a\otimes p)\oplus(b\otimes p).\end{aligned}$$

Therefore, all nine properties hold true, and the set of monic polynomials with all real zeros is a vector space under the operations $\oplus$ and $\otimes$.

□

Now we may analyze the relationship

$$(r_1, r_2, \ldots, r_n) \to \prod_{i=1}^{n}(x-r_i)$$

by thinking of the set of roots of a polynomial as a vector. Therefore, every point in $\mathbb{R}^n$ is associated with a monic polynomial with all real zeros. Since we are interested only in monic polynomials with all real zeros in the interval $[-1,1]$, we may narrow our focus to points in $[-1,1]^n$. This is a compact, convex set. We may now think of norms and other functions that measure some property of a monic polynomial with all real zeros as functions of the form

$$f : [-1,1]^n \to \mathbb{R}.$$

Then, according to the theory of convexity, if we can show that any of these functions are convex, they must achieve a maximum value at one of the corners of $[-1,1]^n$. These corners will correspond to Bernstein polynomials. This idea is still relatively undeveloped. However, it offers a new way of thinking about polynomials, and it provides us with a potentially very powerful tool. Further, if we can use this idea successfully, then we will have a very beautiful and concise explanation of why these Bernstein polynomials possess such maximality.

6 Conclusion

As we have endeavored to show, the geometry of polynomials is a dynamic area of mathematics. There remain unsolved conjectures and new, interesting questions to be asked. Because polynomials are fundamental building blocks for many types of functions, continued pursuit of deeper understanding of them is important. Analyzing how root location influences a polynomial's properties is one way to further this understanding and to continue to expand the field of the geometry of polynomials. We have seen how polynomial root dragging demonstrates intuitive reasons why the Bernstein polynomials frequently arise as functions that maximize certain properties of polynomial functions with zeros in the interval $[-1,1]$: as these functions have their roots in the most extreme locations possible, it makes sense that these results often follow. Further, we have seen some of the wide variations possible among all polynomial functions, whether in measuring their supremum norm, L_1 norm, span of the roots of derivatives, distance from the centroid to the nearest critical number, and more.

Our research has also generated new questions for us to continue to pursue. Some of our ideas for the future include investigating the sensitivity of critical numbers with respect to root motion and trying to quantify how the total change of the critical numbers' locations is distributed among them individually. Also, we aspire to develop fully the idea of using convexity as a tool to provide a more general theory that explains which properties are maximized by Bernstein polynomials. We may also explore other possible benefits that arise from thinking of monic polynomials with all real zeros as a vector space.

The geometry of polynomials is a beautiful and interesting field of mathematics that continues to provide rich problems for study, and it also helps to explain one of the most fundamental objects in mathematics.

References

B. Anderson, Where the Inflection Points of a Polynomial May Lie. *Math Mag*. 70:1 (1997), 32-39.

B. Anderson, Polynomial Root Dragging. *Am. Math. Monthly*. 100:9 (1993), 864-865.

P. Andrews, Where Not to Find the Critical Points of a Polynomial. *Am Math. Monthly*. (1995), 155-158.

M. Boelkins, J. From and S. Kolins, Polynomial Root Squeezing. (2008).

M. Boelkins, J. Miller and B. Vugteveen, From Chebyshev to Bernstein: A Tour of Polynomials Small and Large. *The College Math. Journal 37:3* (200), 194-204.

W. Cheney and D. Kincaid, *Numerical Analysis*, 2nd ed., Brooks/Cole, 1996.

C. Frayer and J. Swenson, Continuous Polynomial Root Dragging, Preprint.

R. Gelca, A Short Proof of a Result on Polynomials *Am. Math. Monthly*. (1993), 936-937.

A. Melman, Bounds on the Zeros of the Derivative of a Polynomial with All Real Zeros. *The Math. Assoc. of Am*. 115 (2008), 145-147.

G.V. Milovanovic, Topics in Polynomials: Extemal Problems, Inequalities, and Zeros. *World Sci. Pub*. (1994).

P. Pawlowski, On the Zeros of a Polynomial and its Derivatives. *Am. Math. Society*. 350:11 (1998), 446 1-4472.

G. Peyser, On the Roots of the Derivative of a Polynomial with Real Roots. *Am. Math. Monthly.* (1967), 1102-1104.

R. Robinson, On the Spans of Derivatives of Polynomials. *Am. Math. Monthly.* (1964), 504-508.

Bl. Sendov, Hausdorff Geometry of Polynomials, *East J. Approx. 7* (2001), 123-178.