

APPROXIMATE ROOTS
AND THE HIDDEN GEOMETRY OF POLYNOMIAL COEFFICIENTS

A Dissertation
Submitted to the Faculty
of
Purdue University
by
Nathan Moses

In Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

May 2020
Purdue University
West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Tzuong-Tsieng Moh, Co-Chair

Department of Mathematics

Dr. William Heinzer, Co-Chair

Department of Mathematics

Dr. Tong Liu

Department of Mathematics

Dr. Louis de Branges

Department of Mathematics

Approved by:

Dr. David Goldberg

Associate Head for Graduate Studies in Mathematics

In dedication to my late Professor Shreeram Shankar Abhyankar, whose socratic approach to math research inspires students across the globe and motivated – as Rama’s bow – this endeavor, which proved so much greater than I had known.

ACKNOWLEDGMENTS

I am grateful to my major professor, Dr. Tzuong-Tsieng Moh, for introducing me to the theory of which this research is a part and for his enduring patience, cheer-leading, and kindness over the years. His commitment to shepherding me through graduation even after his retirement is above what was asked of him.

I would like to thank Professors William Heinzer, Tong Liu, and Louis de Branges for serving on my committee. In particular, I recognize Dr. Heinzer for accepting the position of committee co-chair and for encouraging me to become a student of Professor Abhyankar my first year at Purdue.

I wouldn't be here without the blindly loving support of my family: Chuck, Kim, Charles, Elaine, Clark, Sharon, Heather, and Brandon.

Lastly, I appreciate Ashley Garla, ever my translator in all things.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	vi
ABBREVIATIONS	vii
ABSTRACT	viii
1 INTRODUCTION	1
1.1 Abhyankar-Moh and the Non-Archimedean Case	1
1.2 Approximate Root Polynomials	2
2 ELEMENTARY SYMMETRIC POLYNOMIALS	5
3 2 CLUSTERS OF n ZEROS	7
4 POWER SUMS AND MOMENTS	10
5 m CLUSTERS OF n ZEROS	13
5.1 Case: $m < n$	14
5.2 Case: $n \leq m$	15
6 CONTINUITY OF COMPLEX SOLUTIONS	17
6.1 Rouché's Theorem	17
6.2 Case: $n = m$	18
7 FINDING THE ERROR	20
8 HIDDEN GEOMETRY IN THE COEFFICIENTS OF POLYNOMIALS	25
REFERENCES	26
A APPROXIMATE ROOT PLOTS	27
VITA	31

LIST OF FIGURES

Figure	Page
1.1 Cluster with Gravitational Center	4
A.1 Chains with $2 \leq m \leq 5$	27
A.2 Chains with $6 \leq m \leq 13$	28
A.3 Ladders with $m = 2, 3, 5, 7, 8, 9, 10, 12$	29
A.4 Boundary Plots	30

ABBREVIATIONS

APR approximate root polynomial

ABSTRACT

Moses, Nathan. Ph.D., Purdue University, May 2020. Approximate Roots and the Hidden Geometry of Polynomial Coefficients. Major Professor: Tzuong-Tsieng Moh.

The algebraic operation of approximate roots provides a geometric approximation of the zeros of a polynomial in the complex plane given conditions on their symmetry. A polynomial of degree n corresponds to a cluster of n zeros in the complex plane. The zero of the n^{th} approximate root polynomial locates the gravitational center of this cluster. When the polynomial is of degree mn , with m clusters of n zeros, the centers of the clusters are no longer identified by the zeros of the n^{th} approximate root polynomial in general. The approximation of the centers can be recovered given assumptions about the symmetric distribution of the zeros within each cluster, and given that $m > n$. Rouché's theorem is used to extend this result to relax some of these conditions. This suggests an insight into the geometry of the distribution of zeros within the complex plane hidden within the coefficients of polynomials.

1. INTRODUCTION

The algebraic operation of approximate roots provides a geometric approximation of the zeros of a polynomial in the complex plane given conditions on their symmetry.

1.1 Abhyankar-Moh and the Non-Archimedean Case

The proof of the Abhyankar-Moh Theorem in [1] uses the theory of approximate roots of polynomials first developed in [2] and published in [3]. In particular, the numerical lemma (3.1) in [1] is born from properties of approximate roots demonstrated by the second fundamental theorem for the concept of approximate roots in [2].

The theory of approximate roots initially considered the coefficient field of the polynomials to be Laurent series of an indeterminate. [2] generalizes the coefficient field to any non-archimedean valued field. In such a case, the zeros of a polynomial are located in a tree of discs with the zeros of approximate roots as centers.

What remains uninvestigated are the corresponding properties of approximate roots in an archimedean valued field. To begin that work, this research considers approximate roots of polynomials with complex coefficients, where the archimedean valuation of \mathbb{C} is the complex modulus.

The geometry in \mathbb{C} is rather different than the non-archimedean case. While we can in both spaces consider discs that contain clusters of the zeros of a polynomial, discs in a non-archimedean valued field have the property that all points strictly within the interior of the disc have the same distance to every point on the boundary of the disc, and therefore every point in the interior has an equal claim as the center. From this, it follows that all triangles in this geometry are isosceles; this is a different world than \mathbb{C} indeed!

1.2 Approximate Root Polynomials

$\sqrt{2}$ is famously not an integer nor a ratio of two integers; it can, however, be approximated to arbitrary precision using decimal notation, as can any expression of the form $\sqrt[n]{c}$, for $n, c \in \mathbb{Z}_+$. That is

$$\sqrt[n]{c} = \sum_{i=0}^{\infty} b_i \left(\frac{1}{10}\right)^i$$

where $b_i \in \mathbb{Z}_+$ and given $\{b_0, \dots, b_k\}$ we can find b_{k+1} , and so on. Applications are often satisfied with approximations to this (generally infinite and irregular) series, achieved by truncating as

$$\sqrt[n]{c} \approx \sum_{i=0}^k b_i \left(\frac{1}{10}\right)^i$$

for $k \in \mathbb{Z}_+$. We could, for instance, insist $k = 0$, in which case $\sqrt[n]{c}$ is approximated by an integer, which might have useful properties. For example, $\sqrt{2} \approx 1$ is not precise enough for some applications, but is perhaps useful in defining the movement of a king piece in chess. In general, we can think about approximation to k decimal places as the equivalent of approximating $\sqrt[n]{10^{kn}c}$ with an integer.

This strategy toward approximate roots of integers naturally extends to approximating roots of polynomials; this is the primary consideration of this document.

In the simplest case, the root of a polynomial is again a polynomial. For instance, $\sqrt{x^2 + 2x + 1} = (x + 1)$. How might we approximate $\sqrt{x^2 + 1}$? Similar to the strategy above, we consider

$$\sqrt{x^2 + 1} = \sum_{i=0}^{\infty} b_i \left(\frac{1}{x}\right)^i$$

where $b_i \in \mathbb{C}[x]$. In approximation, we set $b_0 = x$, $b_1 = -\frac{1}{2}$, etc., and we have

$$\sqrt{x^2 + 1} = x - \frac{1}{2}x^{-1} + \dots$$

We can confirm that our approximation is successful thus far by noting that we have minimized the degree of $(x^2 + 1 - (x - \frac{1}{2}x^{-1})^2)$. Truncating this approximation, we can achieve a polynomial approximation, in particular

$$\sqrt{x^2 + 1} \approx x$$

Let us formalize this concept.

Suppose $f \in K[z]$. For $n \mid \deg(f)$, we define an n^{th} approximate root polynomial (abbreviated as ARP) of f to be $g \in K[z]$ such that

$$\deg(f - g^n) < \deg(f) - \deg(g). \quad (1.1)$$

Note that if $\deg(g^n) > \deg(f)$ then $\deg(f - g^n) = \deg(g^n) > \deg(f)$; if $\deg(g^n) < \deg(f)$ then $\deg(f - g^n) = \deg(f)$. Hence the inequality in our definition implies that $\deg(g^n) = \deg(f)$.

Now, if $\deg(f) = n$, i.e.

$$f(z) = \prod_{i=1}^n (z - \delta_i)$$

So,

$$f(z) = z^n \left(1 - \left(\sum_{i=1}^n \delta_i \right) z^{-1} + \dots \right)$$

We quickly see that

$$g(z) = z - \frac{\sum_{i=1}^n \delta_i}{n} \implies g^n(z) = z^n \left(1 - \left(\sum_{i=1}^n \delta_i \right) z^{-1} + \dots \right)$$

So $\deg(f - g^n) \leq n - 2$ and g is an n^{th} approximate root polynomial of f . We note that, for given f and suitable n , g is unique.

Conveniently, any polynomial f can be identified with its zeros in the complex plane, of which there are $\deg(f)$ according to the Fundamental Theorem of Algebra. Therefore, algebraic operations on f can be reconsidered as geometric operations on points in the complex plane. Let us take special notice that in the case of f as above, the zero of g is the arithmetic mean of the zeros of f . Thus, the algebraic operation of calculating an n^{th} approximate root polynomial corresponds to a geometric operation of locating the gravitational center of the cluster of n zeros of f .

In this case, we have found the zero of the n^{th} ARP at the gravitational center of the roots of f . In [2], where K is non-archimedean, any value in the interior of a disc can be identified as a center. In the case of $K = \mathbb{C}$, small perturbations from the center cannot be identified with the center, and therefore we ought to expect those

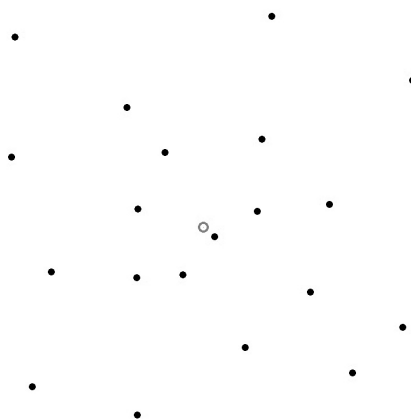


Figure 1.1. Cluster with Gravitational Center

deviations from centers of clusters to accrue. Therefore, the spirit of this project is one of sufficiently bounding the accumulating error in our polynomial approximations that we might preserve the useful properties elaborated in the non-archimedean case [2].

The n^{th} *approximate root polynomial* is achieved in a straightforward way when $\deg(f) = n$, but generally how well does an n^{th} ARP of f geometrically approximate the zeros of f ?

2. ELEMENTARY SYMMETRIC POLYNOMIALS

Before considering *approximate root polynomials* more generally, we introduce notation and obtain some results concerning elementary symmetric polynomials. This simplifies the work to follow, as the coefficients of a polynomial are elementary symmetric polynomials in its zeros.

Consider $x = (x_1, \dots, x_n)$.

Let the k -th elementary symmetric polynomial be defined as

$$e_k(x) = \sum_{1 \leq i_1 < \dots < i_k \leq n} x_{i_1} \dots x_{i_k}$$

for $k > 0$. Let $e_0(x) = 1$.

Then we have the following property:

$$\prod_{i=1}^n (z - x_i) = \sum_{j=0}^n (-1)^j e_j(x) z^{n-j} \quad (2.1)$$

Note that this property holds for any arbitrary n and therefore e_k can be considered as a function of any positive integer number of variables.

Let $x_i = \sigma + \delta_i$ and $\delta = (\delta_1, \dots, \delta_n)$. Then we have the following expansion of elementary symmetric polynomials of binomials:

$$e_k(x) = \sum_{j=0}^k \binom{n-j}{k-j} e_j(\delta) \sigma^{k-j} \quad (2.2)$$

Note that (2.1) above is a special case of (2.2) when $k = n$. Yet we can plug (2.2) into (2.1) to get:

$$\prod_{i=1}^n (z - x_i) = \sum_{j=0}^n (-1)^j \left(\sum_{k=0}^j \binom{n-k}{j-k} e_k(\delta) \sigma^{j-k} \right) z^{n-j} \quad (2.3)$$

Now suppose we have $\omega = \{\omega_1, \dots, \omega_n\}$ so that the ω_i are the n distinct n^{th} -roots of unity, that is:

$$\prod_{i=1}^n (z - \omega_i) = z^n - 1 \quad (2.4)$$

Then by comparing (2.1) and (2.4), we see:

$$e_j(\omega) = \begin{cases} 1 & j = 0 \\ 0 & 0 < j < n \\ (-1)^{n-1} & j = n \end{cases} \quad (2.5)$$

If the x_i are vertices of a regular polygon in the complex plane, they are of the form:

$$x_i = \sigma + r\omega_i$$

where σ and r are constants. Then with $\delta = (r\omega_1, \dots, r\omega_n)$,

$$e_j(\delta) = e_j(\omega)r^j \quad (2.6)$$

and by (2.2), (2.6), and (2.5):

$$\begin{aligned} e_k(x) &= \sum_{j=0}^k \binom{n-j}{k-j} e_j(\delta) \sigma^{k-j} = \sum_{j=0}^k \binom{n-j}{k-j} e_j(\omega) r^j \sigma^{k-j} \\ &= \binom{n}{k} \sigma^k + \binom{0}{k-n} (-1)^{n-1} r^n \sigma^{k-n} = \begin{cases} \binom{n}{k} \sigma^k & 0 \leq k < n \\ \sigma^n + (-1)^{n-1} r^n & k = n \end{cases} \end{aligned} \quad (2.7)$$

Thus, we see how elementary symmetric polynomials yield simplified results when our points are distributed as vertices of regular polygons or, in other words, symmetrically placed on a circle in the complex plane.

3. 2 CLUSTERS OF n ZEROS

Only assuming that we can divide the zeros of a polynomial into two clusters of equal counts, say of n points, we can take its n^{th} *approximate root polynomial*. What can we say about the zeros of the n^{th} ARP?

That is, suppose $f \in K[z]$ is of the form

$$f(z) = \prod_{i=1}^n (z - \sigma_1 - \delta_{1,i})(z - \sigma_2 - \delta_{2,i}).$$

Let $\delta_1 = (\delta_{1,1}, \dots, \delta_{1,n})$ and $\delta_2 = (\delta_{2,1}, \dots, \delta_{2,n})$. By (2.3),

$$\begin{aligned} \prod_{i=1}^n (z - \sigma_1 - \delta_{1,i}) &= \sum_{j=0}^n (-1)^j \left(\sum_{k=0}^j \binom{n-k}{j-k} e_k(\delta_1) \sigma_1^{j-k} \right) z^{n-j} \\ &= z^n \left(1 - \left(\binom{n}{1} e_0(\delta_1) \sigma_1 + \binom{n-1}{0} e_1(\delta_1) \right) z^{-1} \right. \\ &\quad \left. + \left(\binom{n}{2} e_0(\delta_1) \sigma_1^2 + \binom{n-1}{1} e_1(\delta_1) \sigma_1 + \binom{n-2}{0} e_2(\delta_1) \right) z^{-2} + \dots \right) \end{aligned}$$

Without loss of generality, we may choose σ_1 so that $e_1(\delta_1) = 0$. So,

$$\prod_{i=1}^n (z - \sigma_1 - \delta_{1,i}) = z^n \left(1 - \left(\binom{n}{1} \sigma_1 \right) z^{-1} + \left(\binom{n}{2} \sigma_1^2 + e_2(\delta_1) \right) z^{-2} + \dots \right)$$

Similarly, we can have

$$\prod_{i=1}^n (z - \sigma_2 - \delta_{2,i}) = z^n \left(1 - \left(\binom{n}{1} \sigma_2 \right) z^{-1} + \left(\binom{n}{2} \sigma_2^2 + e_2(\delta_2) \right) z^{-2} + \dots \right)$$

So

$$\begin{aligned} f(z) &= z^{2n} \left(1 - \left(\binom{n}{1} \sigma_1 + \binom{n}{1} \sigma_2 \right) z^{-1} \right. \\ &\quad \left. + \left(\binom{n}{2} \sigma_1^2 + e_2(\delta_1) + \binom{n}{2} \sigma_2^2 + e_2(\delta_2) + \binom{n}{1} \binom{n}{1} \sigma_1 \sigma_2 \right) z^{-2} + \dots \right) \end{aligned}$$

Let

$$g(z) = z^2 \left(1 - (\sigma_1 + \sigma_2)z^{-1} + \left(\sigma_1\sigma_2 - \frac{e_2(\delta_1) + e_2(\delta_2)}{n} \right) z^{-2} \right)$$

Then

$$\begin{aligned} g(z)^n &= z^{2n} \left(1 - \binom{n}{1} (\sigma_1 + \sigma_2) z^{-1} \right. \\ &\quad \left. + \left(\binom{n}{2} (\sigma_1 + \sigma_2)^2 + \binom{n}{1} \left(\sigma_1\sigma_2 - \frac{e_2(\delta_1) + e_2(\delta_2)}{n} \right) \right) z^{-2} + \dots \right) \end{aligned}$$

Comparing the coefficients of z in $f(z)$ and $g(z)^n$, some algebra shows that the highest three coefficients agree, so $\deg(f - g^n) < 2n - 2$. Hence, $g(z)$ is our n^{th} *approximate root polynomial* of $f(z)$.

Now the zeros of $g(z)$ can be found by the quadratic formula, which locates them as

$$\begin{aligned} z &= \frac{\sigma_1 + \sigma_2}{2} \pm \sqrt{\left(\frac{\sigma_1 + \sigma_2}{2} \right)^2 - \sigma_1\sigma_2 + \frac{e_2(\delta_1) + e_2(\delta_2)}{n}} \\ &= \frac{\sigma_1 + \sigma_2}{2} \pm \sqrt{\left(\frac{\sigma_1 - \sigma_2}{2} \right)^2 + \frac{e_2(\delta_1) + e_2(\delta_2)}{n}} \end{aligned}$$

If $\frac{e_2(\delta_1) + e_2(\delta_2)}{n} = 0$, then the zeros of $g(z)$ are σ_1 and σ_2 , the centers of our two clusters of zeros of $f(z)$. Since $h(z) = \sqrt{z}$ is continuous, we can regard $\frac{e_2(\delta_1) + e_2(\delta_2)}{n}$ as an error term that contributes to the zeros of our n^{th} *approximate root polynomial* deviating from those centers.

Since

$$2e_2(\delta_1) + \sum_{i=1}^n \delta_{1,i}^2 = e_1(\delta_1)^2 = 0$$

We have $e_2(\delta_1) = -\frac{1}{2} \sum_{i=1}^n \delta_{1,i}^2$ and similarly $e_2(\delta_2) = -\frac{1}{2} \sum_{i=1}^n \delta_{2,i}^2$

So bounding the sum of the squared deviations of the zeros of $f(z)$ from σ_1 and σ_2 can ensure that the zeros of $g(z)$ are arbitrarily close to these cluster centers.

But what does this mean geometrically? For instance, by reparametrization, we can set $\sigma_1 + \sigma_2 = 0$, so $\sigma_1 = -\sigma_2 = \sigma$ and

$$z = \pm \sqrt{-\sigma_1\sigma_2 + \frac{e_2(\delta_1) + e_2(\delta_2)}{n}} = \pm \sqrt{\sigma^2 - \frac{\sum_{i=1}^n \delta_{1,i}^2 + \sum_{i=1}^n \delta_{2,i}^2}{2n}}$$

whose magnitude is the geometric mean of the magnitudes of $\sigma \pm \sqrt{\varepsilon}$ where $\varepsilon = \frac{\sum_{i=1}^n \delta_{1,i}^2 + \sum_{i=1}^n \delta_{2,i}^2}{2n}$, the average of the squares of the $\delta_{j,i}$. The geometric meaning of the average of the $\delta_{j,i}$ is transparent: in terms of physics, this is our center of mass; in terms of statistics, this is our mean. What is the geometric meaning of ε ?

4. POWER SUMS AND MOMENTS

Let the k^{th} *power sum* be defined as

$$p_k(x) = \sum_{i=1}^n x_i^k$$

Then our work in the previous section can be rewritten as

$$2e_2(\delta_i) + p_2(\delta_i) = e_1(\delta_i)^2 = 0$$

This case is generalized by Newton-Girard as

$$ke_k(x) = \sum_{j=1}^k (-1)^{j-1} e_{k-j}(x) p_j(x) \tag{4.1}$$

or alternatively

$$p_k(x) = (-1)^{k-1} ke_k(x) + \sum_{j=1}^{k-1} (-1)^{k-1+j} e_{k-j}(x) p_j(x) \tag{4.2}$$

Since the coefficients of polynomials are elementary symmetric expressions of the zeros, these formulae give us a way to relate the coefficients of our polynomial to power sums of the zeros.

In finding g , our n^{th} *approximate root polynomial* of f , we seek a polynomial that, when raised to the n^{th} power, agrees in the first $d + 1$ coefficients of highest degree, where $d = \text{deg}(g)$. This is a reasonable expectation, since f and hence g can be considered monic, and therefore g has d coefficients to be determined. Because the n^{th} ARP only concerns itself with the first $d + 1$ coefficients, g is not so much approximating just f , but a whole family \mathcal{F} of polynomials, who agree in the first $d + 1$ coefficients. Notably, both f and g^n are members of \mathcal{F} . In this way, the n^{th} *approximate root polynomial* can be considered as a map from the nd dimensional space of monic polynomials of degree nd to the d dimensional space of polynomials

of degree d , where \mathcal{F} is the fiber over g . As \mathcal{F} has $(n - 1)d$ coefficients that are undetermined by g , we realize that much of the information in f was lost in the process of taking the n^{th} ARP. What information has been preserved?

We saw in Chapter 1 that, when $\deg(f) = n$, $\deg(g) = 1$, and our n^{th} *approximate root polynomial* has only one zero. It chooses this zero to be the gravitational center of the zeros of f ; since g approximates all members of \mathcal{F} equally, it follows that the zero of g is the gravitational center of all the members of \mathcal{F} . We now have a new geometric definition of \mathcal{F} : all the polynomials whose roots have the same center as the roots of f .

In Chapter 3, $\deg(g) = 2$, so geometrically g has two zeros with which to approximate f . By reparametrization, we centered f at the origin; notably, that centered g as well. Hence some of the information in g is the center of f , but with two points we expect that g can do better.

By (4.1), with $k = 1, 2, 3$

$$\begin{aligned} e_1(x) &= p_1(x) \\ 2e_2(x) &= e_1(x)p_1(x) - p_2(x) \\ 3e_3(x) &= e_2(x)p_1(x) - e_1(x)p_2(x) + p_3(x) \end{aligned}$$

Since we can always reparametrize so that $e_1(x) = p_1(x) = 0$, we can force

$$\begin{aligned} e_2(x) &= -\frac{p_2(x)}{2} \\ e_3(x) &= \frac{p_3(x)}{3} \end{aligned}$$

So we see that the first few coefficients of a centered f are directly proportional to the power sums of its zeros. The first power sum has much geometric intuition; in physics, it is the center of mass; in statistics, it is the mean; in both, it is called the first moment. In fact, $p_k(x)$ corresponds to the k^{th} moment: for $k = 2$, physics calls this rotational inertia and statistics calls this variance; for higher k , we might have

names like skew and kurtosis, but our names and intuition for the geometric properties they measure quickly run out. (This situation is complicated by the fact that we are operating in the complex plane. This means that squares, and thus second moments, can be negative – or, more generally, complex – while statistics doesn't allow, for instance, negative variance).

With the zeros of g given as $x = (x_1, \dots, x_d)$, let us consider g^n . Since we can assume f is centered at the origin, e_1 of the zeros of f is 0, and the same goes for g^n . e_2 of the zeros of g^n is

$$\binom{n}{1} e_2(x) + \binom{n}{2} e_1(x)^2 = n e_2(x)$$

since $e_1(x) = 0$. In this case, since we know e_2 and p_2 are directly proportional, we know that the zeros of g have $\frac{1}{n}$ the second moment of that of the zeros of f .

We could go further, for higher degrees of g and correspondingly to higher moments, but the author's geometric intuition in this area depends on scientific experience that is inadequate. To quote the translated end of Riemann's *Habilitation Dissertation* [4]:

This path leads out into the domain of another science, into the realm of physics, into which the nature of this present occasion forbids us to penetrate.

5. m CLUSTERS OF n ZEROS

Suppose we have m clusters of n zeros of a polynomial $f(z)$. Then Chapter 3 becomes but the case $m = 2$. We could, naturally, proceed to the cases of three and four clusters, but we soon encounter limitations. Note that in Chapter 3, upon discovering $g(z)$, we used the quadratic formula to locate its zeros. While there are (increasingly cumbersome and therefore relatively opaque) formulae for cubic and quartic polynomials, the Abel-Ruffini Theorem indicates that this strategy proves impossible for higher degrees. Since locating the zeros of an n^{th} *approximate root polynomial* of $f(z)$ when $\deg(f) = mn$ means solving for the zeros of a m^{th} degree polynomial, we must alter our strategy in the general case of m clusters of n zeros.

We realize however, from our experience in Chapter 3, that we can expect the error in the zeros of an n^{th} *approximate root polynomial* of $f(z)$ locating the centers of the clusters of the zeros of $f(z)$ to be in terms of elementary symmetric polynomials in the deviations within the clusters from the center of the cluster, $\delta_{j,i}$. In Chapter 2, we found in (2.4) – (2.7) that almost all of the elementary symmetric polynomials of the $\delta_{j,i}$ vanish when the clusters of the zeros are arranged as the vertices of a regular polygon. We leverage this assumption in pursuit of the results of this section.

Suppose $f \in K[z]$ is of the form

$$f(z) = \prod_{j=1}^m \prod_{i=1}^n (z - \sigma_j - \delta_{j,i})$$

Let

$$\begin{aligned} x_{j,i} &= \sigma_j + \delta_{j,i} \\ x_j &= (x_{j,1}, \dots, x_{j,n}) \\ \sigma_j^* &= (\sigma_j, \dots, \sigma_j) \\ \delta_j &= (\delta_{j,1}, \dots, \delta_{j,n}) \\ f_j(z) &= \prod_{i=1}^n (z - x_{j,i}) \end{aligned}$$

Then

$$f(z) = \prod_{j=1}^m \prod_{i=1}^n (z - x_{j,i}) = \prod_{j=1}^m f_j(z)$$

We use (2.2) to get

$$e_k(x_j) = \sum_{l=0}^k \binom{n-l}{k-l} e_l(\delta_j) \sigma_j^{k-l}$$

If we assume the zeros of each $f_j(z)$ are vertices of a regular polygon, we may suppose $\delta_{j,i} = r_j \omega_i$ where r_j is a constant, and $\{\omega_i\}$ are the n distinct n^{th} -roots of unity. Then by (2.5) and (2.6)

$$e_l(\delta_j) = \begin{cases} 1 & l = 0 \\ 0 & 0 < l < n \\ (-1)^{n-1} r_j^l & l = n \end{cases}$$

5.1 Case: $m < n$

Then for $0 \leq k \leq m$

$$e_k(x_j) = \binom{n}{k} \sigma_j^k$$

and

$$f_j(z) = \prod_{i=1}^n (z - x_{j,i}) = \sum_{k=0}^n (-1)^k e_k(x_j) z^{n-k}$$

whose highest $m + 1$ degree terms agree with the highest $m + 1$ degree terms of

$$\prod_{i=1}^n (z - \sigma_j) = \sum_{k=0}^n (-1)^k e_k(\sigma_j^*) z^{n-k} = \sum_{k=0}^n (-1)^k \binom{n}{k} \sigma_j^k z^{n-k}$$

As the highest $m + 1$ degree terms of $f(z)$ are determined by the highest $m + 1$ degree terms of the $f_j(z)$, the highest $m + 1$ degree terms of $f(z)$ must agree with the highest $m + 1$ degree terms of

$$f^*(z) = \prod_{j=1}^m \prod_{i=1}^n (z - \sigma_j)$$

Now, the n^{th} *approximate root polynomial* of $f^*(z)$ is clearly

$$g(z) = \prod_{j=1}^m (z - \sigma_j)$$

as $f^*(z) - g(z)^n = 0$.

Since $g(z)$ is determined by the first $m + 1$ terms of $f^*(z)$ and $\deg(g) = m$, $g(z)$ is evidently also an n^{th} ARP of $f(z)$. As in Chapter 4, we can think of f and f^* as both being in the fiber above g .

As $g(z)$ is in factored form, we see that the zeros of $g(z)$ are exactly the centers of the clusters, σ_j . Hence, when each cluster is symmetrical and the number of zeros in each cluster exceeds the number of clusters, we precisely locate the centers with the zeros of the n^{th} *approximate root polynomial*. We could think of this in the sense that each cluster has enough influence on $g(z)$ that the n^{th} ARP ignores the pull of other clusters in determining each cluster's center.

5.2 Case: $n \leq m$

Then for $0 \leq k \leq n$

$$e_k(x_j) = \begin{cases} \binom{n}{k} \sigma_j^k & k < n \\ \sigma_j^n + (-1)^{n-1} r_j^n & k = n \end{cases}$$

So

$$\begin{aligned} f_j(z) &= \prod_{i=1}^n (z - x_{j,i}) = \sum_{k=0}^n (-1)^k e_k(x_j) z^{n-k} = -r_j^n + \sum_{k=0}^n (-1)^k \binom{n}{k} \sigma_j^k z^{n-k} \\ &= -r_j^n + \sum_{k=0}^n (-1)^k e_k(\sigma_j^*) z^{n-k} = -r_j^n + \prod_{i=1}^n (z - \sigma_j) \end{aligned}$$

And

$$f(z) = \prod_{j=1}^m \left(-r_j^n + \prod_{i=1}^n (z - \sigma_j) \right)$$

Note that the highest n degree terms of $f_j(z)$ agree with the highest n degree terms of $\prod_{i=1}^n (z - \sigma_j)$. So, the highest n degree terms of $f(z)$ agree with the highest n degree terms of

$$f^*(z) = \prod_{j=1}^m \prod_{i=1}^n (z - \sigma_j)$$

The next highest degree term of $f(z)$, that of degree $nm - n$, has a coefficient that differs from the corresponding coefficient of $f^*(z)$ by $\sum_{j=1}^m r_j^n$. In general, lower degree terms have coefficient differences in the form of a sum of a product of some r_j^n and some elementary symmetric polynomials of all the σ_i , save σ_j .

As the n^{th} *approximate root polynomial* of $f^*(z)$ is clearly

$$g^*(z) = \prod_{j=1}^m (z - \sigma_j)$$

and $\deg(g^*) = m$, the $m + 1$ coefficients of $g^*(z)$ are determined by the highest $m + 1$ degree terms of $f^*(z)$. In particular, the highest n degree terms of $g^*(z)$ are determined by the highest n degree terms of $f(z)$. Unfortunately, $m + 1 > n$, so $g^*(z)$ is not generally the n^{th} *approximate root polynomial* of $f(z)$ (unless, say, all $r_j = 0$, the trivial case). Hence, the r_j constitute an error term; bounding the radius of our clusters ever tighter yields a better and better approximation of the cluster's center. We may think of this as each cluster being more tight-knit or, alternatively, more distinct. Can we make the effect of this error term more precise?

6. CONTINUITY OF COMPLEX SOLUTIONS

Just as the coefficients of a polynomial are elementary symmetric polynomials of the polynomial's zeros, and therefore the coefficients vary continuously according to the zeros, so too do the zeros depend continuously on the coefficients.

We can see this clearly using Rouché's Theorem.

6.1 Rouché's Theorem

Let f and g be analytic in a simply connected domain U . Let C be a simple closed contour in U . If $|g(z)| < |f(z)|$ for every z on C , then the functions $f(z)$ and $f(z) + g(z)$ have the same number of zeros, counting multiplicities, inside C .

Proof. Let $h_t(z) = f(z) + tg(z)$. Then, by the argument principle, the number of zeros of $h_t(z)$ in the interior of C is

$$\frac{1}{2\pi i} \oint_C \frac{h_t'(z)}{h_t(z)} dz$$

Note that

$$|h_t(z)| = |f(z) + tg(z)| \geq |f(z)| - t|g(z)| \geq |f(z)| - |g(z)| > 0$$

So $h_t(z)$ is a homotopy between $f(z)$ and $f(z) + g(z)$ such that the number of zeros of $h_t(z)$ in the interior of C is a continuous, integer-valued function of $t \in [0, 1]$, and hence constant.

Say $f(z)$ is a polynomial with m zeros strictly within $\varepsilon > 0$ of some z_0 . We know that $g(z) = f(z) + \delta$ with

$$|\delta| < \min_{|z-z_0|=\varepsilon} |f(z)|$$

then for all z on the boundary of the disc of radius ε around z_0

$$|g(z) - f(z)| = |\delta| < |f(z)|$$

and by Rouché's Theorem, $g(z)$ has exactly m roots within ε of z_0 as well.

(It would be nice if we had a multivariate generalization for Rouché's Theorem. For instance, suppose we have a system of equations $g_j(\varepsilon_1, \dots, \varepsilon_n) = 0$ for $1 \leq j \leq n$, where each g_j is a polynomial of degree j in the variables ε_i . Can we use a generalization of Rouché's Theorem to determine a bound on the solution of the new system of equations $\tilde{g}_j = 0$ where $\tilde{g}_j = g_j + \delta_j$?)

In Chapter 5, we saw that the difference between coefficients of $f(z)$ and $f^*(z)$ were in terms of products involving r_j^k . This inequality indicates that we can bound the zeros of our polynomial arbitrarily close to the centers of our clusters by sufficiently bounding r_j , i.e. the radii of the clusters.

Let us make better use of this technique by examining a specific case.

6.2 Case: $n = m$

Then as before, $g^*(z) = \prod_{j=1}^m (z - \sigma_j)$ and the n^{th} approximate root polynomial of $f(z)$ is

$$g(z) = \frac{\sum_{j=1}^m -r_j^n}{n} + \prod_{j=1}^m (z - \sigma_j) = \frac{\sum_{j=1}^m -r_j^n}{n} + g^*(z)$$

Then for $z \in C_{R,j}$, the circle of radius R around σ_j , so $z = \sigma_j + R\zeta$, for $|\zeta| = 1$, we want

$$|g(z) - g^*(z)| = \frac{\left| \sum_{j=0}^m r_j^n \right|}{n} < \min_{z \in C_{R,j}} |g^*(z)| = \min_{|\zeta|=1} |R \prod_{\substack{i=1 \\ i \neq j}}^m (\sigma_j - \sigma_i + R\zeta)|$$

If

$$R < |\sigma_j - \sigma_i + R\zeta| < |\sigma_j - \sigma_i| - R$$

That is

$$R < \frac{|\sigma_j - \sigma_i|}{2}$$

Then

$$\frac{\left| \sum_{j=1}^m r_j^n \right|}{n} < R^m$$

is a sufficient condition for $|g(z) - g^*(z)| < |g^*(z)|$ on $C_{R,j}$ and hence for $g(z)$ to have a zero within $C_{R,j}$.

If we say that r is the maximum of the r_j , then $\frac{\left| \sum_{j=0}^m r_j^n \right|}{n} < r^n < R^m$ becomes $r < R$, since $n = m$. Thus, so long as our clusters are sufficiently separated ($r < \frac{|\sigma_j - \sigma_i|}{2}$ for all i, j), then the zeros of our approximate root is within r of the corresponding center. Specifically, if for all j , $r_j = r$, then we have that the zeros of the approximate roots are each located within the disc associated with each cluster.

A final note about continuity of complex solutions: As small deviations of the zeros of a polynomial produce similarly small deviations in coefficients (and vice versa) we realize that we might, in practice, be able to relax our assumptions about the symmetry in the zeros of our polynomial. Thus, our work here represents an ideal circumstance that likely bears useful results in wider applications.

7. FINDING THE ERROR

Let

$$f(z) = \prod_{i=1}^m \prod_{j=1}^n (z - \sigma_i - \delta_{i,j})$$

An n^{th} *approximate root polynomial* (APR) of f is a polynomial

$$g(z) = \prod_{i=1}^m (z - \sigma_i - \varepsilon_i)$$

such that

$$g(z)^n = \prod_{i=1}^m (z - \sigma_i - \varepsilon_i)^n = \prod_{i=1}^m \prod_{j=1}^n (z - \sigma_i - \varepsilon_i)$$

whose coefficients of the highest $m + 1$ degree z terms agree with the corresponding z terms of f .

Note that if $x_i = x$ for all i , we have $e_k(x_i) = \binom{n}{k} x^k$.

$$g(z)^n = \sum_{j=0}^n \left(\sum_{\substack{j_1 + \dots + j_m = j \\ 0 \leq k_i \leq j_i}} \prod_{i=1}^m \binom{n - k_i}{j_i - k_i} e_{k_i}(\varepsilon_i) \sigma_i^{j_i - k_i} \right) z^{n-j} \quad (7.1)$$

We also can choose the c_i so that $e_1(\delta_i) = 0$.

Using (7.1), we look at the coefficients of z corresponding to $1 \leq j \leq m$. (Note both polynomials are monic, so $j = 0$ is trivial.)

Case $m = 1$

$$(j = 1) \implies \binom{n}{1} \varepsilon_1 = 0$$

Case $m = 2, n = 1$

$$(j = 1) \implies \varepsilon_1 + \varepsilon_2 = 0$$

$$(j = 2) \implies \varepsilon_2 \sigma_1 + \varepsilon_1 \sigma_2 + \varepsilon_1 \varepsilon_2 = 0$$

Case $m = 2, n = 2$

$$(j = 1) \implies \binom{2}{1}\varepsilon_1 + \binom{2}{1}\varepsilon_2 = 0$$

$$(j = 2) \implies \binom{2}{1}\varepsilon_1\sigma_1 + \varepsilon_1^2 + \binom{2}{1}\binom{2}{1}\varepsilon_2\sigma_1 + \binom{2}{1}\binom{2}{1}\varepsilon_1\sigma_2 + \binom{2}{1}\binom{2}{1}\varepsilon_1\varepsilon_2$$

$$+ \binom{2}{1}\varepsilon_2\sigma_2 + \varepsilon_2^2 = e_2(\delta_1) + e_2(\delta_2)$$

Case $m = 2, n > 2$

$$(j = 1) \implies \binom{n}{1}\varepsilon_1 + \binom{n}{1}\varepsilon_2 = 0$$

$$(j = 2) \implies \binom{n-1}{1}\binom{n}{1}\varepsilon_1\sigma_1 + \binom{n}{2}\varepsilon_1^2 + \binom{n}{1}\binom{n}{1}\varepsilon_2\sigma_1 + \binom{n}{1}\binom{n}{1}\varepsilon_1\sigma_2$$

$$+ \binom{n}{1}\binom{n}{1}\varepsilon_1\varepsilon_2 + \binom{n-1}{1}\binom{n}{1}\varepsilon_2\sigma_2 + \binom{n}{2}\varepsilon_2^2$$

$$= e_2(\delta_1) + e_2(\delta_2)$$

Case $m = 3, n = 1$

$$(j = 1) \implies \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0$$

$$(j = 2) \implies \varepsilon_2\sigma_1 + \varepsilon_3\sigma_1 + \varepsilon_1\sigma_2 + \varepsilon_1\varepsilon_2 + \varepsilon_1\sigma_3$$

$$+ \varepsilon_1\varepsilon_3 + \varepsilon_3\sigma_2 + \varepsilon_2\sigma_3 + \varepsilon_2\varepsilon_3 = 0$$

$$(j = 3) \implies \varepsilon_3\sigma_1\sigma_2 + \varepsilon_2\sigma_1\sigma_3 + \varepsilon_2\varepsilon_3\sigma_1 + \varepsilon_1\sigma_2\sigma_3 + \varepsilon_1\varepsilon_3\sigma_2$$

$$+ \varepsilon_1\varepsilon_2\sigma_3 + \varepsilon_1\varepsilon_2\varepsilon_3 = 0$$

Case $m = 3, n = 2$

$$(j = 1) \implies \binom{2}{1}\varepsilon_1 + \binom{2}{1}\varepsilon_2 + \binom{2}{1}\varepsilon_3 = 0$$

$$(j = 2) \implies \binom{2}{1}\varepsilon_1\sigma_1 + \varepsilon_1^2 + \binom{2}{1}\binom{2}{1}\varepsilon_2\sigma_1 + \binom{2}{1}\binom{2}{1}\varepsilon_3\sigma_1 + \binom{2}{1}\binom{2}{1}\varepsilon_1\sigma_2$$

$$+ \binom{2}{1}\binom{2}{1}\varepsilon_1\varepsilon_2 + \binom{2}{1}\binom{2}{1}\varepsilon_1\sigma_3 + \binom{2}{1}\binom{2}{1}\varepsilon_1\varepsilon_3 + \binom{2}{1}\varepsilon_2\sigma_2 + \varepsilon_2^2$$

$$\begin{aligned}
& + \binom{2}{1} \binom{2}{1} \varepsilon_3 \sigma_2 + \binom{2}{1} \binom{2}{1} \varepsilon_2 \sigma_3 + \binom{2}{1} \binom{2}{1} \varepsilon_2 \varepsilon_3 + \binom{2}{1} \varepsilon_3 \sigma_3 + \varepsilon_3^2 \\
& \quad = e_2(\delta_1) + e_2(\delta_2) + e_2(\delta_3) \\
(j=3) \implies & \binom{2}{1} \varepsilon_2 \sigma_1^2 + \binom{2}{1} \varepsilon_3 \sigma_1^2 + \binom{2}{1} \binom{2}{1} \varepsilon_1 \sigma_1 \sigma_2 + \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_2 \sigma_1 + \binom{2}{1} \binom{2}{1} \varepsilon_1 \sigma_1 \sigma_3 \\
& + \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_3 \sigma_1 + \binom{2}{1} \varepsilon_1^2 \sigma_2 + \binom{2}{1} \varepsilon_1^2 \varepsilon_2 + \binom{2}{1} \varepsilon_1^2 \sigma_3 + \binom{2}{1} \varepsilon_1^2 \varepsilon_3 \\
& + \binom{2}{1} \binom{2}{1} \varepsilon_2 \sigma_1 \sigma_2 + \binom{2}{1} \varepsilon_2^2 \sigma_1 + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_3 \sigma_1 \sigma_2 + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_2 \sigma_1 \sigma_3 \\
& + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_2 \varepsilon_3 \sigma_1 + \binom{2}{1} \binom{2}{1} \varepsilon_3 \sigma_1 \sigma_3 + \binom{2}{1} \varepsilon_3^2 \sigma_1 + \binom{2}{1} \varepsilon_1 \sigma_2^2 + \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_2 \sigma_2 \\
& + \binom{2}{1} \varepsilon_1 \varepsilon_2^2 + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_1 \sigma_2 \sigma_3 + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_3 \sigma_2 + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_2 \sigma_3 \\
& + \binom{2}{1} \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_2 \varepsilon_3 + \binom{2}{1} \varepsilon_1 \sigma_3^2 + \binom{2}{1} \binom{2}{1} \varepsilon_1 \varepsilon_3 \sigma_3 + \binom{2}{1} \varepsilon_1 \varepsilon_3^2 + \binom{2}{1} \varepsilon_3 \sigma_2^2 \\
& + \binom{2}{1} \binom{2}{1} \varepsilon_2 \sigma_2 \sigma_3 + \binom{2}{1} \binom{2}{1} \varepsilon_2 \varepsilon_3 \sigma_2 + \binom{2}{1} \varepsilon_2^2 \sigma_3 + \binom{2}{1} \varepsilon_2^2 \varepsilon_3 + \binom{2}{1} \binom{2}{1} \varepsilon_3 \sigma_2 \sigma_3 \\
& \quad + \binom{2}{1} \varepsilon_3^2 \sigma_2 + \binom{2}{1} \varepsilon_2 \sigma_3^2 + \binom{2}{1} \binom{2}{1} \varepsilon_2 \varepsilon_3 \sigma_3 + \binom{2}{1} \varepsilon_2 \varepsilon_3^2 \\
& = \binom{2}{1} e_2(\delta_1) \sigma_2 + \binom{2}{1} e_2(\delta_1) \sigma_3 + \binom{2}{1} e_2(\delta_2) \sigma_1 + \binom{2}{1} e_2(\delta_3) \sigma_1 + \binom{2}{1} e_2(\delta_2) \sigma_3 + \binom{2}{1} e_2(\delta_3) \sigma_2
\end{aligned}$$

Case $m = 3, n = 3$

$$\begin{aligned}
(j=1) \implies & \binom{3}{1} \varepsilon_1 + \binom{3}{1} \varepsilon_2 + \binom{3}{1} \varepsilon_3 = 0 \\
(j=2) \implies & \binom{2}{1} \binom{3}{1} \varepsilon_1 \sigma_1 + \binom{3}{2} \varepsilon_1^2 + \binom{3}{1} \binom{3}{1} \varepsilon_2 \sigma_1 + \binom{3}{1} \binom{3}{1} \varepsilon_3 \sigma_1 + \binom{3}{1} \binom{3}{1} \varepsilon_1 \sigma_2 \\
& + \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_2 + \binom{3}{1} \binom{3}{1} \varepsilon_1 \sigma_3 + \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_3 + \binom{2}{1} \binom{3}{1} \varepsilon_2 \sigma_2 + \binom{3}{2} \varepsilon_2^2 \\
& + \binom{3}{1} \binom{3}{1} \varepsilon_3 \sigma_2 + \binom{3}{1} \binom{3}{1} \varepsilon_2 \sigma_3 + \binom{3}{1} \binom{3}{1} \varepsilon_2 \varepsilon_3 + \binom{2}{1} \binom{3}{1} \varepsilon_3 \sigma_3 + \binom{3}{2} \varepsilon_3^2 \\
& \quad = e_2(\delta_1) + e_2(\delta_2) + e_2(\delta_3) \\
(j=3) \implies & \binom{3}{1} \varepsilon_1 \sigma_1^2 + \binom{3}{2} \varepsilon_1^2 \sigma_1 + \varepsilon_1^3 + \binom{3}{2} \binom{3}{1} \varepsilon_2 \sigma_1^2 + \binom{3}{2} \binom{3}{1} \varepsilon_3 \sigma_1^2 \\
& + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \sigma_1 \sigma_2 + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_2 \sigma_1 + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \sigma_1 \sigma_3
\end{aligned}$$

$$\begin{aligned}
& + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_3 \sigma_1 + \binom{3}{1} \binom{3}{2} \varepsilon_1^2 \sigma_2 + \binom{3}{2} \binom{3}{1} \varepsilon_1^2 \varepsilon_2 + \binom{3}{1} \binom{3}{2} \varepsilon_1^2 \sigma_3 + \binom{3}{2} \binom{3}{1} \varepsilon_1^2 \varepsilon_3 \\
& + \binom{3}{1} \binom{2}{1} \binom{3}{1} \varepsilon_2 \sigma_1 \sigma_2 + \binom{3}{1} \binom{3}{2} \varepsilon_2^2 \sigma_1 + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_3 \sigma_1 \sigma_2 + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_2 \sigma_1 \sigma_3 \\
& \quad + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_2 \varepsilon_3 \sigma_1 + \binom{3}{1} \binom{2}{1} \binom{3}{1} \varepsilon_3 \sigma_1 \sigma_3 + \binom{3}{1} \binom{3}{2} \varepsilon_3^2 \sigma_1 + \binom{3}{2} \binom{3}{1} \varepsilon_1 \sigma_2^2 \\
& + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_2 \sigma_2 + \binom{3}{1} \binom{3}{2} \varepsilon_1 \varepsilon_2^2 + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \sigma_2 \sigma_3 + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_3 \sigma_2 \\
& + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_2 \sigma_3 + \binom{3}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_2 \varepsilon_3 + \binom{3}{2} \binom{3}{1} \varepsilon_1 \sigma_3^2 + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_1 \varepsilon_3 \sigma_3 \\
& \quad + \binom{3}{1} \binom{3}{2} \varepsilon_1 \varepsilon_3^2 + \binom{3}{1} \varepsilon_2 \sigma_2^2 + \binom{3}{2} \varepsilon_2^2 \sigma_2 + \varepsilon_2^3 + \binom{3}{2} \binom{3}{1} \varepsilon_3 \sigma_2^2 \\
& + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_2 \sigma_2 \sigma_3 + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_2 \varepsilon_3 \sigma_2 + \binom{3}{1} \binom{3}{2} \varepsilon_2^2 \sigma_3 + \binom{3}{2} \binom{3}{1} \varepsilon_2^2 \varepsilon_3 \\
& + \binom{3}{1} \binom{2}{1} \binom{3}{1} \varepsilon_3 \sigma_2 \sigma_3 + \binom{3}{1} \binom{3}{2} \varepsilon_3^2 \sigma_2 + \binom{3}{2} \binom{3}{1} \varepsilon_2 \sigma_3^2 + \binom{2}{1} \binom{3}{1} \binom{3}{1} \varepsilon_2 \varepsilon_3 \sigma_3 \\
& \quad + \binom{3}{1} \binom{3}{2} \varepsilon_2 \varepsilon_3^2 + \binom{3}{1} \varepsilon_3 \sigma_3^2 + \binom{3}{2} \varepsilon_3^2 \sigma_3 + \varepsilon_3^3 \\
& = e_2(\delta_1)\sigma_1 + e_3(\delta_1) + \binom{3}{1} e_2(\delta_1)\sigma_2 + \binom{3}{1} e_2(\delta_1)\sigma_3 + \binom{3}{1} e_2(\delta_2)\sigma_1 + \binom{3}{1} e_2(\delta_3)\sigma_1 \\
& \quad + e_2(\delta_2)\sigma_2 + e_3(\delta_2) + \binom{3}{1} e_2(\delta_2)\sigma_3 + \binom{3}{1} e_2(\delta_3)\sigma_2 + e_2(\delta_3)\sigma_3 + e_3(\delta_3)
\end{aligned}$$

Case $m = 3, n > 3$

$$\begin{aligned}
(j = 1) & \implies \binom{n}{1} \varepsilon_1 + \binom{n}{1} \varepsilon_2 + \binom{n}{1} \varepsilon_3 = 0 \\
(j = 2) & \implies \binom{n-1}{1} \binom{n}{1} \varepsilon_1 \sigma_1 + \binom{n}{2} \varepsilon_1^2 + \binom{n}{1} \binom{n}{1} \varepsilon_2 \sigma_1 + \binom{n}{1} \binom{n}{1} \varepsilon_3 \sigma_1 + \binom{n}{1} \binom{n}{1} \varepsilon_1 \sigma_2 \\
& + \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_2 + \binom{n}{1} \binom{n}{1} \varepsilon_1 \sigma_3 + \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_3 + \binom{n-1}{1} \binom{n}{1} \varepsilon_2 \sigma_2 + \binom{n}{2} \varepsilon_2^2 \\
& + \binom{n}{1} \binom{n}{1} \varepsilon_3 \sigma_2 + \binom{n}{1} \binom{n}{1} \varepsilon_2 \sigma_3 + \binom{n}{1} \binom{n}{1} \varepsilon_2 \varepsilon_3 + \binom{n-1}{1} \binom{n}{1} \varepsilon_3 \sigma_3 + \binom{n}{2} \varepsilon_3^2 \\
& = e_2(\delta_1) + e_2(\delta_2) + e_2(\delta_3) \\
(j = 3) & \implies \binom{n-1}{2} \binom{n}{1} \varepsilon_1 \sigma_1^2 + \binom{n-2}{1} \binom{n}{2} \varepsilon_1^2 \sigma_1 + \binom{n}{3} \varepsilon_1^3 + \binom{n}{2} \binom{n}{1} \varepsilon_2 \sigma_1^2 + \binom{n}{2} \binom{n}{1} \varepsilon_3 \sigma_1^2
\end{aligned}$$

$$\begin{aligned}
& + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \sigma_1 \sigma_2 + \binom{n-1}{1} \binom{n}{1} \varepsilon_1 \binom{n}{1} \varepsilon_2 \sigma_1 + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \sigma_1 \sigma_3 \\
& + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_3 \sigma_1 + \binom{n}{1} \binom{n}{2} \varepsilon_1^2 \sigma_2 + \binom{n}{2} \binom{n}{1} \varepsilon_1^2 \varepsilon_2 + \binom{n}{1} \binom{n}{2} \varepsilon_1^2 \sigma_3 + \binom{n}{2} \binom{n}{1} \varepsilon_1^2 \varepsilon_3 \\
& + \binom{n}{1} \binom{n-1}{1} \binom{n}{1} \varepsilon_2 \sigma_1 \sigma_2 + \binom{n}{1} \binom{n}{2} \varepsilon_2^2 \sigma_1 + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_3 \sigma_1 \sigma_2 + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_2 \sigma_1 \sigma_3 \\
& + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_2 \varepsilon_3 \sigma_1 + \binom{n}{1} \binom{n-1}{1} \binom{n}{1} \varepsilon_3 \sigma_1 \sigma_3 + \binom{n}{1} \binom{n}{2} \varepsilon_3^2 \sigma_1 + \binom{n}{2} \binom{n}{1} \varepsilon_1 \sigma_2^2 \\
& + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_2 \sigma_2 + \binom{n}{1} \binom{n}{2} \varepsilon_1 \varepsilon_2^2 + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \sigma_2 \sigma_3 + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_3 \sigma_2 \\
& + \binom{n}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_2 \sigma_3 + \binom{n}{1} \varepsilon_1 \binom{n}{1} \varepsilon_2 \binom{n}{1} \varepsilon_3 + \binom{n}{2} \binom{n}{1} \varepsilon_1 \sigma_3^2 + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_1 \varepsilon_3 \sigma_3 \\
& + \binom{n}{1} \binom{n}{2} \varepsilon_1 \varepsilon_3^2 + \binom{n-1}{2} \binom{n}{1} \varepsilon_2 \sigma_2^2 + \binom{n-2}{1} \binom{n}{2} \varepsilon_2^2 \sigma_2 + \binom{n}{3} \varepsilon_2^3 + \binom{n}{2} \binom{n}{1} \varepsilon_3 \sigma_2^2 \\
& + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_2 \sigma_2 \sigma_3 + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_2 \varepsilon_3 \sigma_2 + \binom{n}{1} \binom{n}{2} \varepsilon_2^2 \sigma_3 + \binom{n}{2} \binom{n}{1} \varepsilon_2^2 \varepsilon_3 \\
& + \binom{n}{1} \binom{n-1}{1} \binom{n}{1} \varepsilon_3 \sigma_2 \sigma_3 + \binom{n}{1} \binom{n}{2} \varepsilon_3^2 \sigma_2 + \binom{n}{2} \binom{n}{1} \varepsilon_2 \sigma_3^2 + \binom{n-1}{1} \binom{n}{1} \binom{n}{1} \varepsilon_2 \varepsilon_3 \sigma_3 \\
& \quad + \binom{n}{1} \binom{n}{2} \varepsilon_2 \varepsilon_3^2 + \binom{n-1}{2} \binom{n}{1} \varepsilon_3 \sigma_3^2 + \binom{n-2}{1} \binom{n}{2} \varepsilon_3^2 \sigma_3 + \binom{n}{3} \varepsilon_3^3 \\
& = \binom{n-2}{1} e_2(\delta_1) \sigma_1 + e_3(\delta_1) + \binom{n}{1} e_2(\delta_1) \sigma_2 + \binom{n}{1} e_2(\delta_1) \sigma_3 + \binom{n}{1} e_2(\delta_2) \sigma_1 + \binom{n}{1} e_2(\delta_3) \sigma_1 \\
& + \binom{n-2}{1} e_2(\delta_2) \sigma_2 + e_3(\delta_2) + \binom{n}{1} e_2(\delta_2) \sigma_3 + \binom{n}{1} e_2(\delta_3) \sigma_2 + \binom{n-2}{1} e_2(\delta_3) \sigma_3 + e_3(\delta_3)
\end{aligned}$$

...

We observe that when, for all i , all $\delta_{i,j}$ are equal to the same δ_i , we have $\varepsilon_i = \delta_i$. This in particular is always the case when $n = 1$. How do we see that from the equations?

This is precisely the situation where our parenthetical desire in the previous section for a multidimensional generalization of Rouché's Theorem would be helpful. Intuitively, small deviations on the right hand side from zero should have small effects on the ε_i .

8. HIDDEN GEOMETRY IN THE COEFFICIENTS OF POLYNOMIALS

Working with ARPs has drawn our attention to the patterns and information within the coefficients of polynomials. We have noted that

$$f(z) = z^n + c$$

geometrically represents a regular n -gon in the complex plane, and

$$f(z) = (z^n + c)z^m = z^{n+m} + cz^m$$

represents the same, but with m points also added at the figure's center.

We know from Rouché's Theorem that any polynomial with nearly the same coefficients as f has nearly the same geometric arrangement of points. Therefore we can think of the n^{th} coefficient after the highest degree polynomial term as carrying information about the n -gon-ness of that polynomial's points. Further, since we know that a polynomial of degree n can represent any arrangement of n points in the complex plane, we can consider that the first $n + 1$ terms of a polynomial determine n points that it generally resembles. For instance, the first 4 terms of a polynomial can be considered as containing information about the triangleness of its roots.

Naturally, as the ARP considers the first coefficients only, an ARP is an n^{th} root approximation of this geometric sketch.

REFERENCES

- [1] S. S. Abhyankar and T. T. Moh. Embeddings of the line in the plane. *J. reine angew. Math.*, 276:148–166, 1975.
- [2] T. T. Moh. On two fundamental theorems for the concept of approximate roots. *J. Math. Soc. Japan*, 34(4):637–652, 1982.
- [3] S. S. Abhyankar and T. T. Moh. Newton-puiseux expansion and generalized tschirnhausen transformation. *J. reine angew. Math.*, 260:47–63, 1973.
- [4] B. Riemann (translated by W. K. Clifford). On the hypotheses which lie at the bases of geometry. *Nature*, 8(183-184):14–17, 1998.

A. APPROXIMATE ROOT PLOTS

This appendix provides visual representations of the zeros of ARP (in blue) and the zeros of the polynomials they approximate (red). All figures were created by software developed exclusively by the author.

Since we found that the ARP located centers of clusters best when $n > m$, we examine the circumstances where we expect the ARP to perform most poorly, i.e. when $n = 2$ and m grows.

At the end, we consider plots of a couple cases of polynomials of small degree (4 and 6 respectively) where the approximate root zero is just on the boundary of the discs containing the clusters of roots.

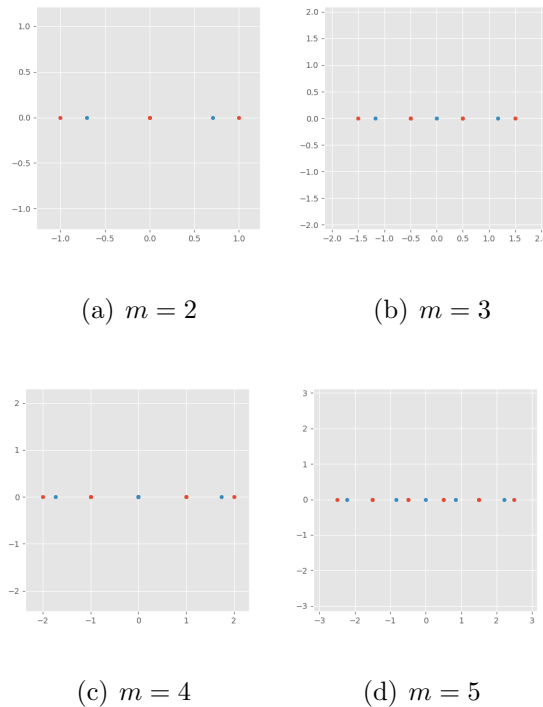
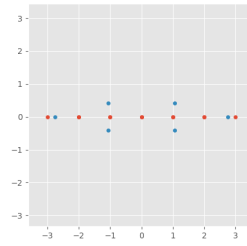
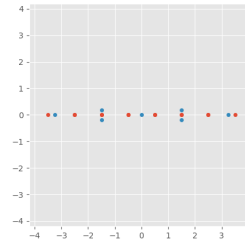
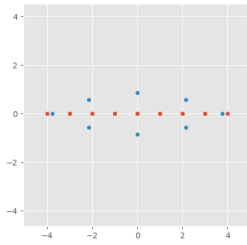
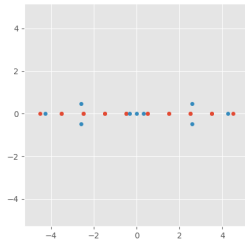
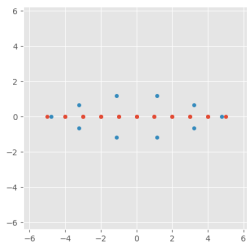
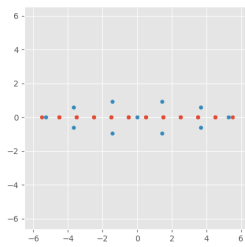
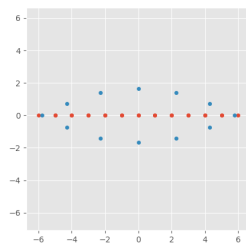
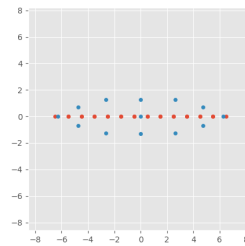
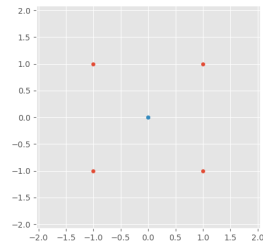
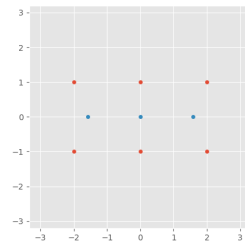
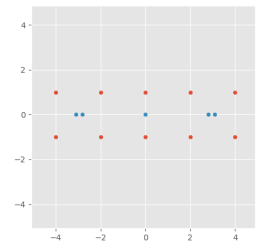
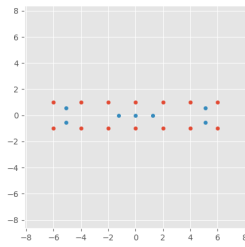
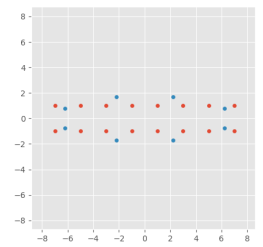
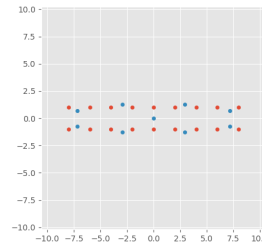
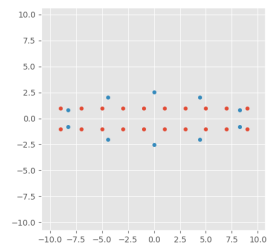
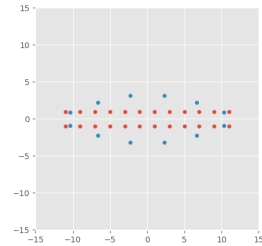
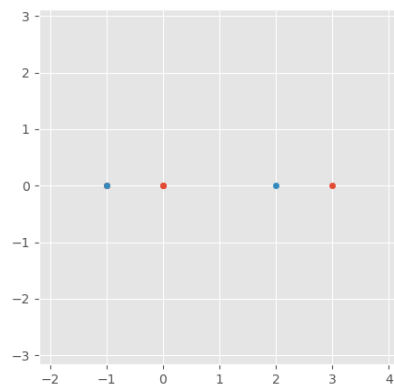


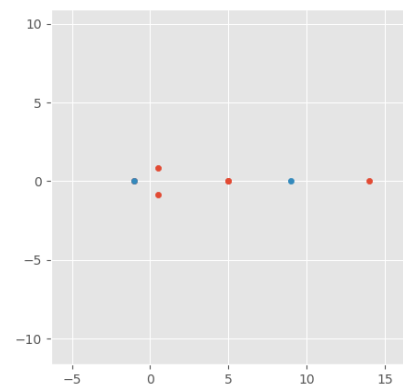
Figure A.1. Chains with $2 \leq m \leq 5$

(a) $m = 6$ (b) $m = 7$ (c) $m = 8$ (d) $m = 9$ (e) $m = 10$ (f) $m = 11$ (g) $m = 12$ (h) $m = 13$ Figure A.2. Chains with $6 \leq m \leq 13$

(a) $m = 2$ (b) $m = 3$ (c) $m = 5$ (d) $m = 7$ (e) $m = 8$ (f) $m = 9$ (g) $m = 10$ (h) $m = 12$ Figure A.3. Ladders with $m = 2, 3, 5, 7, 8, 9, 10, 12$



(a) Degree 4



(b) Degree 6

Figure A.4. Boundary Plots

VITA

Nathan Moses was born on July 23, 1989 in Dayton, Ohio. He received a B.A. in Philosophy and a B.S. in Mathematics in 2011 from Miami University in Oxford, Ohio. During his undergraduate studies, he worked on *An efficient algorithm for the computation of average mutual information* with Dr. Robin Thomas and presented it at the 2009 Society for Mathematical Psychology meeting in Amsterdam (it was published in 2014 in the Journal of Mathematical Psychology). He attended the University of Hyderabad in the Spring of 2010. He began graduate studies at Purdue University in August 2011. The Spring and Summer of 2012, he worked with S. S. Abhyankar on two papers, *Explicit Multivariable Meromorphic Series* and *Galois Theory Of Meromorphic Series*, the former of which was approved for publication the day after Dr. Abhyankar's passing.