Roots of Random Polynomials

Tamique de Brito, Tony Wang, Yang Yan

November 2019

1 Introduction

The roots of polynomials are interesting mathematical objects. The fundamental theorem of algebra says that given $a_0, \ldots, a_n \in \mathbb{C}$, there exist $r_1, \ldots, r_n \in \mathbb{C}$ unique up to permutation such that

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n = (z - r_1)(z - r_2) \cdots (z - r_n).$$

In other words, the roots of a complex polynomial are uniquely determined by its coefficients. While it is easy to find the coefficients of a polynomial given its roots, finding the roots from coefficients is a much harder problem.

In this paper we investigate a related problem: analyzing the distribution of the roots of a polynomial with random coefficients. Our results center around the following phenomenon: for high degree random polynomials, the distribution a random root is close to that of the uniform distribution on the unit circle.

We begin, in Section 2, by looking at some simple arguments that hint at why the unit circle concentration behavior is true. In Section 3, we hypothesize and prove that when our random polynomials have coefficients which are chosen from circularly symmetric distributions, the joint root distribution is angularly uniform. In Section 4, we build up some deterministic bounds on the roots of a general polynomial through the use of a polynomial's companion matrix, preparing us to prove our central result.

Finally, in Section 5, we give a proof of our central result: the distribution of a random root is close to that of a uniform distribution on the unit circle. We prove this in the case of a complex polynomial of the form $A_0 + \cdots + A_{n-1} z^{n-1} + z^n$ where $A_0, \ldots, A_{n-1}$ are i.i.d. complex random variables. We find that under weak conditions on the coefficient distribution, as $n \to \infty$, a random root is near the unit circle. Furthermore, if the A_i are all bounded by a constant, then the arguments of a random root approaches the uniform distribution.

To conclude, we also propose several conjectures in Section 6, motivated by numerical evidence, regarding further properties of the distributions of roots.

1.1 Probability Notation

We use $\mathbf{P}(A)$ to denote the probability of an event A. We use $\mathbf{E}[X]$ to denote the expectation of a random variable X. Random variables are generally denoted using capital letters.

2 Building some intuition

In this section we build up some intuition using simple arguments as to why we expect high degree polynomials to have roots that lie near the unit circle.

2.1 Roots of random binomials of high degree

We will investigate the simple case where the random polynomial is in the form $z^n + A$, where A is a random variable, and find that the distribution of the magnitude of the roots concentrates around 1 as the degree goes to infinity. Specifically, we will look at the probability distribution of the roots of the random polynomial $z^n + A$ as $n \to \infty$ where A is any random variable that takes on the value 0 with probability zero.

[This result follows primarily from the fact that for any positive a, $\lim_{n\to\infty} a^n = \infty$ if $a > 1$ and $\lim_{n\to\infty} a^n = \infty$ if $a < 1$.]

Note that if A is represented in complex exponential form $|A|e^{i\mathrm{Arg}(A)}$ the roots of $z^n + A$ are $\xi_i = |A|^{1/n} e^{i\mathrm{Arg}(A)i/n}$, $i = 0, \ldots, n$. Then the magnitude of each root is $|A|^{1/n}$. Now write for each n, ϵ, write $E_{n,\epsilon}$ to denote the event that $1 - \epsilon \leq |\xi_i| \leq 1 + \epsilon$ $i = 1, \ldots, n$ for the polynomial $z^n + A$. Then, our current goal is to show that for all $\epsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}(E_{n,\epsilon}) = 1. \tag{2.1}$$

Now, for a given $a < 1$ the probability that the magnitude of the roots will be less than a is $\mathbf{P}(|\xi_i| < a) = \mathbf{P}(|A|^{1/n} < a) = \mathbf{P}(|A| < a^n)$. Since $\lim_{n\to\infty} a^n = 0$, then

$$\lim_{n\to\infty} \mathbf{P}(|A| < a^n) = 0.$$

Similarly, for $b > 1$, the probability that the magnitude of the roots will be greater than b is $\mathbf{P}(|\xi_i| > b) = \mathbf{P}(|A|^{1/n} > b) = 1 - \mathbf{P}(|A| < b^n)$ and since $\lim_{n\to\infty} b^n = \infty$,

$$\lim_{n\to\infty} 1 - \mathbf{P}(|A| < b^n) = 1 - 1 = 0.$$

Then, $\mathbf{P}(E_{n,\epsilon}) = 1 - \mathbf{P}(E_{n,\epsilon}^C) \geq 1 - (\mathbf{P}(|\xi_i| < 1 - \epsilon) + \mathbf{P}(|\xi_i| > 1 + \epsilon))$, so that

$$\lim_{n\to\infty} \mathbf{P}(E_{n,\epsilon}) \geq \lim_{n\to\infty} (1 - (\mathbf{P}(|\xi_i| < 1 - \epsilon) + \mathbf{P}(|\xi_i| > 1 + \epsilon))) = 1,$$

and therefore,

$$\lim_{n\to\infty} \mathbf{P}(E_{n,\epsilon}) = 1.$$

Note that this reasoning can be applied to binomials in the form $z^n + Az^m$, where m is fixed and $n > m$. Simply factoring out z^m gives the form $z^m(z^{n-m} + A)$ and as $n \to \infty$, $n - m \to \infty$ as well, and in the limit, the m zero roots contributed by the z^m factor are dominated by the $n - m$ roots contributed by the $z^{n-m} + A$, meaning that as $n \to \infty$, the proportion of roots contributed by $z^{n-m} + A$ approaches 1. Note that in this case, it is not true that $1 - \epsilon \leq |\xi_i| \leq 1 + \epsilon$ since some roots of $z^m(z^{n-m} + A)$ are zero. However, as $n \to \infty$, for a root ξ selected uniformly at random, with high probability it will hold that $1 - \epsilon \leq |\xi| \leq 1 + \epsilon$. As a final generalization, note that this can be applied to polynomials in the form $z^n + Az^{m_n}$, where the degree m_n of the non-leading term Az^{m_n} changes with n. In this case, the reasoning goes through if $m_n = o(n)$.

2.2 Convergence of expectation of roots in log space

Here we will prove that the expected value of the logarithm of a root randomly selected from a random n-th degree polynomial approaches zero as n approaches infinity given a that the expectation of the logarithm of the magnitude of the constant coefficient is finite.

Specifically, let $P(z) = A_0 + \cdots + A_{n-1}z^{n-1} + z^n$ be a random polynomial where A_i are i.i.d and R_i be the n roots of this polynomial. Let R be the random variable defined by uniformly selecting one of $R_1, \ldots, R_n$. Now we will show that if $\mathbf{E} \log |A_0|$ is finite, then

$$\lim_{n\to\infty} \mathbf{E}[\log(|R|)] = 0. \tag{2.2}$$

Note that $\mathbf{E}[\log(|R|)] = \frac{1}{n}\sum \mathbf{E}[\log(|R_i|)]$. The fact that $A_0 = \prod R_i$ implies $\log(|A_0|) = \sum \log(|R_i|) = n\log(|R|)$. Then $\mathbf{E}[\log(|R|)] = \frac{1}{n}\mathbf{E}[\log(|A_0|)]$. Therefore, if $\mathbf{E}[\log(|A_0|)]$ has a finite value, then $\log(|R|)$ converges to zero.

Note that this does not rely on any coefficients other than A_0, so this result holds with arbitrary A_i for $i > 0$. Furthermore, this can be generalized to coefficients whose distribution changes with n. Specifically, if $A_0^{(n)}$ is the constant coefficient associated with the n-th degree polynomial, then $\mathbf{E}\log|A_0^{(n)}| = o(n)$ is a sufficient condition for the result to hold.

3 Circularly symmetric coefficients yield angularly uniform roots

From the cases analyzed earlier, we have gained some intuition about why the roots of random high degree polynomials have magnitudes close to 1. We now shift gears and look at the behavior of the argument of the roots. First, we formalize the random variables. We will use $\mathrm{Arg}(z)$ to denote the argument of complex z, which takes values in $[0, 2\pi)$.

Definition 3.1. *Let $A_0, \ldots, A_{n-1}$ be complex-valued random variables. We denote by $R_1, \ldots, R_n$ the complex-valued random variables that satisfy*

$$A_0 + A_1 z + \cdots + A_{n-1} z^{n-1} + z^n = (z - R_1) \cdots (z - R_n), \tag{3.1}$$

$$(|R_1|, \operatorname{Arg}(R_1)) \geq \cdots \geq (|R_n|, \operatorname{Arg}(R_1)), \tag{3.2}$$

where (3.2) is specified using the lexicographic order of $\mathbb{R} \times \mathbb{R}$ and $\operatorname{Arg} : \mathbb{C} \to [0, 2\pi)$. We denote by R the random variable that is a uniform random choice of one of $R_1, \ldots, R_n$.

For example, the lexicographic ordering of the R_i is defined such that complex number $1 + i = (\sqrt{2}, \frac{\pi}{4})$ would be ordered after complex number $i = (1, \frac{\pi}{2})$.

3.1 Numerical results

We ran three simulations for polynomials of degree $n = 10$ and $n = 25$ where the A_i are i.i.d. complex Gaussians with mean 0. For $n = 10$, we sampled from complex Gaussians of both standard deviation 1 and 4. For $n = 100$, we sampled from complex Gaussians with standard deviation 1 only. For each of the three simulations, we sampled 5000 random polynomials. For each sampled polynomial, we numerically computed the roots via the polynomial solver in the `numpy` package in Python. In Figure 3.1 we plot, for both both values of n, the roots in the complex plane as a heatmap, as well as arguments of the roots as a binned histogram.

From our simulations, we suspect that a the roots of random polynomials with A_i as complex Gaussians with mean 0 have arguments which are uniformly distributed in $[0, 2\pi)$. We prove a stronger version of this claim below.

3.2 Proof of angular uniformity of joint root distribution

Recall first the definition of circular symmetry in the complex numbers.

Definition 3.2. *A distribution D in the complex numbers is circularly symmetric if D is identical under rotation about point 0. That is, $D = e^{i\phi} D$ for any angle ϕ,*

We now show that the distribution of the arguments of the roots are circularly symmetric for all circularly symmetric A_i, as opposed to the more specific case of Gaussian A_i as in the simulation.

Theorem 3.3. *Let $A_1, \ldots, A_n$ be independent circularly symmetric distribution. Then, the polynomial $p(z) = A_0 + A_1 z + \ldots + A_{n-1} z^{n-1} + z^n$ has joint root distribution R (as defined in 3.1) which is also circularly symmetric.*

Proof. Consider the random polynomial $q(z) = B_0 + B_1 z + \ldots + B_{n-1} z^{n-1} + z^n$ whose joint root distribution is $e^{i\phi} R$ for any ϕ. From Vieta's formulas, we see that the B_j must be defined as follows:

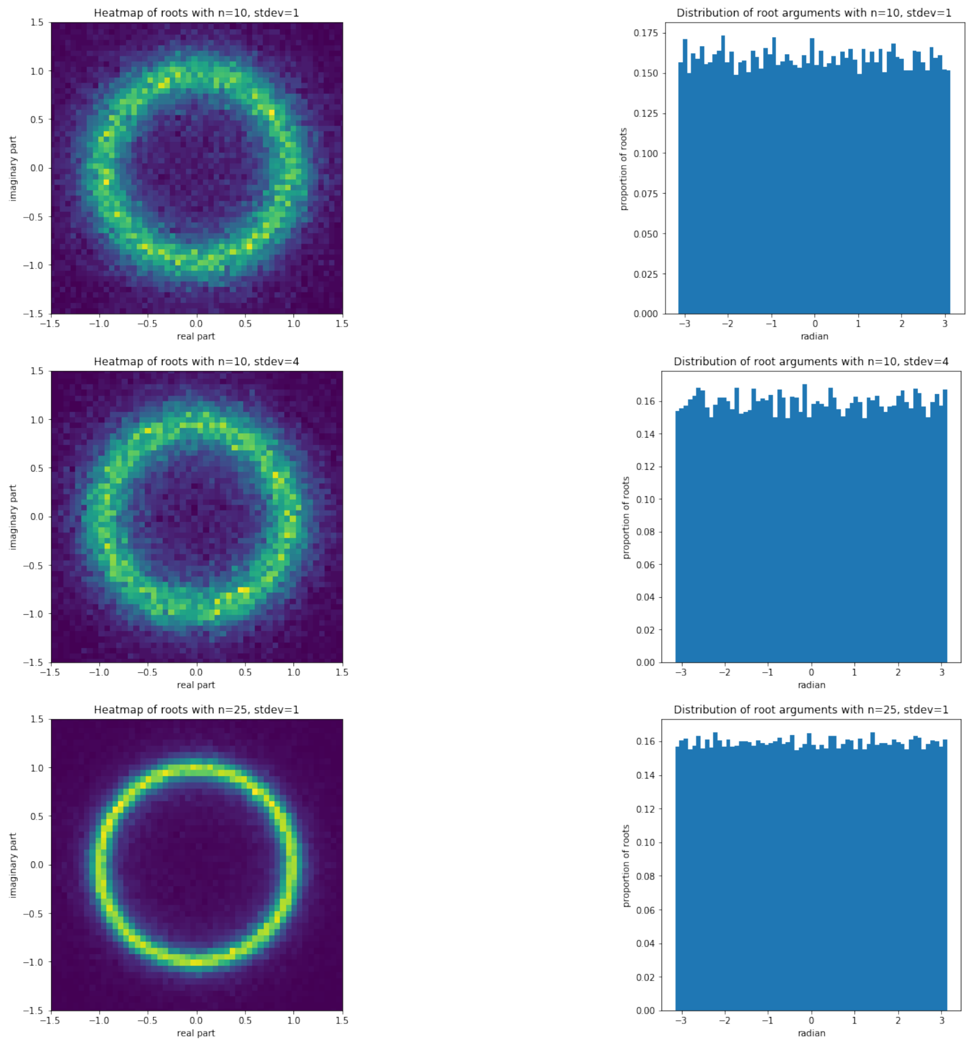

Figure 3.1: Heatmap and histogram of the roots and the arguments of the roots. In the heatmaps, yellow colors denote high concentrations of roots, while purple colors denote lower concentrations. It appears that for complex Gaussian A_i with mean 0, neither the standard deviation of the A_i nor the degree n of the polynomial matter - the arguments of the roots remain uniformly distributed over $[0, 2\pi)$ regardless.

$$B_j = e^{i\phi(n-j)} A_j.$$

However, as all A_j are circularly symmetric, we have that $B_j = A_j$ for all j. We know that there is a unique mapping from a polynomial's coefficients to its roots. Thus, because the coefficients of $p(z)$ and $q(z)$ are sampled from the same distribution, we expect the joint root distribution $e^{i\phi}R$ for any ϕ to be identical to R, the joint root distribution for $p(z)$. By definition, then, the joint root distribution R is circularly symmetric. $\square$

We now show that all distributions over the complex numbers which are circularly symmetric must also be angularly uniform. That is, the distribution of the arguments of each of the points in the distribution is uniform over $[0, 2\pi)$.

Corollary 3.3.1. *The joint root distributions R of polynomials chosen from circularly symmetric A_i are not only circularly symmetric but also angularly uniform.*

Proof. We show that any complex distribution R which is circularly symmetric must be angularly uniform as well. Divide the distribution R into q sections $S_1, \dots, S_q$, each spanning an arc of $1/q$ radians around 0, such that none of the S_i overlap. Then the measure of each of these sections is equivalent. That is,

$$\mathbf{P}(R \in S_i) = \mathbf{P}(R \in S_j)$$

for all i and j, as R is circularly symmetric. Then, $\mathbf{P}(R \in S_i) = \frac{1}{q}$ for all i. Let S'_p be the union of p of the non-overlapping S_i. We must also have, then, that $\mathbf{P}(R \in S'_p) = \frac{p}{q}$. That is, $\mathbf{P}(R \in S)$ is simply the measure of S for any set S with rational measure.

It is well-known that a distribution R which is defined as such for all rational sets S must also be defined for all irrational sets T as simply the measure of S by taking the limit of the rational sets near the irrational set T. R must be angularly uniform, then, as its measure on any subset of angles is simply the measure of the subset. □

We have shown now that the joint root distribution of the polynomials with coefficients chosen from circularly symmetric distributions must be angularly uniform.

4 Deterministic results on the singular values of companion matrices

Before proceeding further, we derive an important inequality involving the roots of a polynomial and its coefficients. This inequality will be used later on to show that the roots of a random high degree polynomial concentrate about the unit circle. We begin by defining for complex polynomial $p(z)$ its *companion matrix* C as follows.

Definition 4.1 (Companion Matrix)**.** *The companion matrix C to a polynomial*

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

is the $n \times n$ matrix

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}. \tag{4.1}$$

The companion matrix of a polynomial $p(z)$ has the property that its characteristic polynomial is precisely the original polynomial $p(z)$. Thus, the eigenvalues of C are the roots of $p(z)$. This is a very powerful transformation of the polynomial. We begin to showcase the power of the companion matrix by deriving a simple bound on the magnitudes of its eigenvalues.

Lemma 4.2. *The maximum magnitude of the eigenvalues r_i is bounded as follows:*

$$\max_i |r_i| \leq 1 + \max_i |a_i|.$$

Proof. Consider any eigenvector $v = (v_1, \ldots, v_n)^T$ of the companion matrix C, scaled such that $\max_i |v_i| = 1$. We then have, for all $i \in [1, n]$:

$$|(Cv)_i| = \sum_{j=1}^{n} |C_{ij} v_j| \leq \sum_{j=1}^{n} |C_{ij}| \leq 1 + \max_i |a_i|.$$

We know the left-hand side, by definition, is equal to $|(r_j v)_i| = |r_j||v_i|$ for some j. We can then write

$$|r_j| \leq \frac{1 + \max_i |a_i|}{\max_i |v_i|}.$$

Observing that $\max_i |v_i| = 1$ gives the desired inequality. □

Knowing the strength of this transformation, we now proceed to further bound the eigenvalues of C and thus the roots of $p(z)$ using the singular values of C. We recall what the singular values of a matrix are.

Definition 4.3 (Singular Values)**.** *The n singular values s_i of a square $n \times n$ matrix M, are the square-roots of the eigenvalues of M^*M, where M^* is the conjugate transpose of M. The singular values are all non-negative real numbers and we order them as $s_1 \geq s_2 \geq \cdots \geq s_n$.*

The singular values of a companion matrix C are characterized by the following lemma.

Lemma 4.4. *Let C be the companion matrix to complex polynomial*

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

Let s_i be the singular values of C. We have that

$$s_1 = \frac{\sqrt{(1 + |a_0|)^2 + \sum_{i=1}^{n-1} |a_i|^2} + \sqrt{(1 - |a_0|)^2 + \sum_{i=1}^{n-1} |a_i|^2}}{2},$$
$$s_2 = s_3 = \cdots = s_{n-1} = 1,$$
$$s_n = \frac{|a_0|}{s_1}.$$

Proof. Consider C^*C:

$$C^*C = \begin{bmatrix} 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \\ -\bar{a}_1 & -\bar{a}_2 & \cdots & -\bar{a}_{n-1} & \sum_{i=1}^n |a_i|^2 \end{bmatrix}. \tag{4.2}$$

The first $n-1$ columns of $C^*C - I$ have zeros everywhere except the last row. So $C^*C - I$ has at most 2 independent columns, a rank of at most 2, and thus, by the rank-nullity theorem[1], a null-space of size at least $n-2$. Now, we recall the well-known result that the algebraic multiplicity of an eigenvalue λ is lower-bounded by the geometric multiplicity of the same eigenvalue λ. In our case, the geometric multiplicity of $\lambda = 1$ is at least $n-2$, so the algebraic multiplicity of $\lambda = 1$ in C^*C must be at least $n-2$. Ignoring the ordering of the s_i for now, we have $s_2 = \cdots = s_{n-1} = 1$.

Now recall that the eigenvalues of a matrix multiply to the determinant of the matrix. We can see that $|\det C| = |a_0|$. Since $\det M^* = \overline{\det M}$. We then have

$$\prod_{i=1}^n s_i^2 = s_1^2 s_n^2 = \det C^*C = |a_0|^2.$$

Furthermore, recall that the eigenvalues of a matrix sum to the trace of the matrix:

$$\sum_{i=1}^n s_i^2 = s_1^2 + (n-2) + s_n^2 = \mathrm{Tr}(C^*C) = (n-1) + \sum_{i=0}^n |a_i|^2.$$

Solving these two equations gives the desired values for s_1 and s_n. □

We now cite a result which bounds the eigenvalues of a matrix with its singular values.

Lemma 4.5 (Weyl's Majorant Inequality [2, Thm 3.1.13])**.** *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and let $A \in \mathbb{C}^{n \times n}$ be a nonsingular matrix with eigenvalues $\lambda_1, \dots, \lambda_n$ and singular values $s_1, \dots, s_n$. Then*

$$\sum_{i=1}^n \varphi(\log|\lambda_i|) \leq \sum_{i=1}^n \varphi(\log s_i).$$

We proceed to bound the eigenvalues of C, and thus the roots of $p(z)$, using our characterization of the singular values as well as Weyl's Majorant Inequality.

Lemma 4.6. *Let $r_1, \dots, r_n$ be the roots of polynomial*

$$p(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n, \ a_0 \neq 0.$$

Then

$$\sum_{i=1}^n \log^2 |r_i| \leq 2\log^2 |a_0| + 3\log^2 \left(1 + \sum_{i=0}^{n-1} |a_i|\right). \tag{4.3}$$

Proof. Let C be the companion matrix to $p(z)$ and let $s_1, \ldots, s_n$ be the singular values of C, ordered as in Lemma 4.4. Since $a_0 \neq 0$, C is nonsingular. This lets us apply Lemma 4.5 with the convex function $\varphi(x) = x^2$. Since the eigenvalues of C are the roots of $p(z)$, we obtain the inequality

$$\sum_{i=1}^{n} \log^2 |r_i| \leq \sum_{i=1}^{n} \log^2 s_i.$$

By Lemma 4.4 we have that

$$\begin{gathered} 1 \leq s_1 \leq \sqrt{(1+|a_0|)^2 + \sum_{i=1}^{n-1} |a_i|^2} \leq 1 + \sum_{i=0}^{n-1} |a_i|, \\ s_2 = s_3 = \cdots = s_{n-1} = 1, \\ s_n = \frac{|a_0|}{s_1}. \end{gathered}$$

This means

$$\begin{aligned} \sum_{i=1}^{n} \log^2 |r_i| &\leq \sum_{i=1}^{n} \log^2 s_i \\ &\leq \log^2 s_1 + \log^2 \frac{|a_0|}{s_1} \\ &= \log^2 s_1 + (\log|a_0| - \log s_1)^2 \\ &\leq 2\log^2 |a_0| + 3\log^2 s_1 \\ &\leq 2\log^2 |a_0| + 3\log^2 \left(1 + \sum_{i=0}^{n-1} |a_i|\right). \end{aligned}$$ □

With this upper-bound on the second moments of the logs of the roots of general polynomials, we may now approach analyzing the distribution of roots for general random polynomials.

5 Roots of random high degree polynomials

We begin our analysis of the roots of random high degree polynomials by proving Theorem 5.2, which gives sufficient conditions for when a random root of a random high degree polynomial has magnitude close to 1. We will then prove Theorem 5.3, which gives sufficient conditions for when a random root of a random high degree polynomial has a uniform angle distribution.

5.1 Roots tend to the unit circle

We now give a proof of the phenomenon that roots of random high degree polynomials tend to have magnitude close to 1. We shall do so in the case of i.i.d. coefficients, but the proof can be extended to the non i.i.d. case as well. The main (and really only) ingredient in our proof is the inequality derived in Lemma 4.2.

In our proof we will make use of the notion of convergence in probability, which we now define.

Definition 5.1 (Convergence in Probability)**.** *Let X_n be a sequence of complex random variables. We say that X_n converges to a random variable X in probability, written as $X_n \xrightarrow{p} X$, when for all $\epsilon > 0$*

$$\lim_{n\to\infty} \mathbf{P}\left(|X_n - X| > \epsilon\right) = 0. \tag{5.1}$$

We are ready to state and prove our result.

Theorem 5.2. *Let $A_0, A_1, \ldots$ be a sequence of i.i.d. complex random variables with finite expectations that satisfy $\mathbf{P}(A_i = 0) = 0$. Let R_n be a random root of the polynomial*

$$p_n(z) = A_0 + A_1 z + \cdots + A_{n-1} z^{n-1} + z^n.$$

Then $|R_n| \xrightarrow{p} 1$.

Proof. We shall show the equivalent statement that $\log |R_n| \xrightarrow{p} 0$.

Let $R_{n,1}, \ldots, R_{n,n}$ denote the roots of $p_n(z)$ and denote

$$\Lambda = \frac{1}{n} \sum_{i=1}^{n} \log^2 |R_{n,i}|.$$

Λ is the empirical second moment of the discrete uniform distribution over $\log |R_{n,1}|$, …, $\log |R_{n,n}|$. We shall show that $\Lambda \xrightarrow{p} 0$, which will imply that $\log |R_n| \xrightarrow{p} 0$.

Lemma 4.6 tells us that

$$\Lambda \leq \frac{1}{n} \log^2 |A_0| + \frac{1}{n} \log^2 \left(1 + \sum_{i=0}^{n-1} |A_i|\right). \tag{5.2}$$

Now

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \mathbf{E}\left[\log^2\left(1 + \sum_{i=0}^{n-1} |A_i|\right)\right] \\
&\leq \lim_{n\to\infty} \frac{1}{n} \mathbf{E}\left[\log^2\left(e + \sum_{i=0}^{n-1} |A_i|\right)\right] \\
&\leq \lim_{n\to\infty} \frac{1}{n} \log^2\left(e + \sum_{i=0}^{n-1} \mathbf{E}\left[|A_i|\right]\right) \qquad \text{(Jensen's inequality)} \\
&\leq \lim_{n\to\infty} \frac{\log^2\left(e + n\mathbf{E}\left[|A_0|\right]\right)}{n} = 0,
\end{aligned}$$

so we have that

$$\frac{1}{n}\log^2\left(1+\sum_{i=0}^{n-1}|A_i|\right)\xrightarrow{p}0. \tag{5.3}$$

Moreover, since $\mathbf{P}(A_0=0)=0$, we also have that

$$\frac{1}{n}\log^2|A_0|\xrightarrow{p}0. \tag{5.4}$$

Since Λ is non-negative, equations (5.2), (5.3), and (5.4) tell us that $\Lambda\xrightarrow{p}0$. □

5.2 Angular uniformity of roots

We have shown that under a mild moment conditions for the coefficients, the magnitude of the roots of a random polynomial tends to be close to 1. We shall now show that under much stronger moment conditions on the coefficients (e.g. i.i.d. coefficients with finite support) the argument of the random root of a random polynomial tends to be uniformly distributed in $[0,2\pi)$.

Theorem 5.3. *Let $A_0, A_1, \ldots$ be a sequence of i.i.d. complex random variables which have finite moments of all orders and satisfy $\mathbf{P}(A_i=0)=0$. Let R_n be a random root of the polynomial*

$$p_n(z)=A_0+A_1z+\cdots+A_{n-1}z^{n-1}+z^n.$$

If for every positive integer w, it holds that

$$\lim_{n\to\infty}\mathbf{E}\left[|R_n^w|\right]=1, \tag{5.5}$$

then for any $0\le a\le b<2\pi$, we have

$$\lim_{n\to\infty}\mathbf{P}\left(\operatorname{Arg}(R_n)\in[a,b]\right)=\frac{b-a}{2\pi}. \tag{5.6}$$

In other words, $\operatorname{Arg}(R_n)$ converges in distribution to the uniform distribution on $[0,2\pi)$.

Proof. To show that equation (5.6) holds for all $0\le a\le b<2\pi$, we will prove the equivalent statement that

$$\lim_{n\to\infty}\mathbf{E}\left[f\left(\operatorname{Arg}(R_n)\right)\right]=\int_0^{2\pi}f(x)\,dx \tag{5.7}$$

for all continuous functions $f:[0,2\pi]\to\mathbb{R}$ with $f(0)=f(2\pi)$.

By Fejer's theorem, for any continuous function $f:[0,2\pi]\to\mathbb{R}$ with $f(0)=f(2\pi)$, there exists a sequence of functions $f_m:[0,2\pi]\to\mathbb{R}$ that converges uniformly to f where

$$f_m(x)=\int_0^{2\pi}f(x)\,dx+\sum_{-m\le w\le m,\,p\ne 0}c_{n,k}e^{iwx}$$

for some constants $c_{n,k}$. We shall show that for any positive integer w,

$$\lim_{n\to\infty} \mathbf{E}\left[\exp\left(iw \operatorname{Arg}\left(R_n\right)\right)\right] = 0, \tag{5.8}$$

which implies that

$$\lim_{n\to\infty} \mathbf{E}\left[f_m\left(\operatorname{Arg}(R_n)\right)\right] = \int_0^{2\pi} f(x)\, dx$$

for all m. This means

$$\begin{aligned}
\limsup_{n\to\infty} \mathbf{E}\left[f\left(\operatorname{Arg}(R_n)\right)\right] &\leq \limsup_{n\to\infty} \mathbf{E}\left[f_m\left(\operatorname{Arg}(R_n)\right) + \|f - f_m\|_\infty\right] \\
&= \int_0^{2\pi} f(x)\, dx + \|f - f_m\|_\infty, \\
\liminf_{n\to\infty} \mathbf{E}\left[f\left(\operatorname{Arg}(R_n)\right)\right] &\leq \limsup_{n\to\infty} \mathbf{E}\left[f_m\left(\operatorname{Arg}(R_n)\right) - \|f - f_m\|_\infty\right] \\
&= \int_0^{2\pi} f(x)\, dx - \|f - f_m\|_\infty,
\end{aligned}$$

which, upon taking the limit as $m \to \infty$, yields equation (5.7).

Thus it remains to show equation (5.8) holds for positive w. To do so we make use of Newton's identities. Letting $R_{n,1}, \ldots, R_{n,n}$ denote the roots of $p_n(z)$, Newton's identities tell us that we can write

$$\sum_{i=1}^{n} R_{n,i}^p = N_p(A_{n-1}, ..., A_{n-p}), \tag{5.9}$$

where N_p is a fixed polynomial (not varying in n) in p unknowns and A_{n-p} is taken to be zero when $n < p$.

Since the A_i have moments of all orders and are i.i.d., the expected value of expression (5.9) is constant for $n \geq p$. Moreover, by Proposition 4.2 the expected value of each individual $R_{n,k}^p$ is also finite. Thus,

$$\lim_{n\to\infty} \mathbf{E}\left[R_n^p\right] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}\left[R_{n,i}^p\right] = \lim_{n\to\infty} \frac{1}{n} \mathbf{E}\left[\sum_{i=1}^{n} R_{n,i}^p\right] = 0. \tag{5.10}$$

We may now show equation (5.8) as follows:

$$\begin{aligned}
&\lim_{n\to\infty} \mathbf{E}\left[\exp\left(ip \operatorname{Arg}\left(R_n\right)\right)\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[R_n^p + (1 - |R_n|^p) \exp\left(ip \operatorname{Arg}\left(R_n\right)\right)\right] \\
&= \lim_{n\to\infty} \mathbf{E}\left[R_n^p\right] + \lim_{n\to\infty} \mathbf{E}\left[(1 - |R_n|^p) \exp\left(ip \operatorname{Arg}\left(R_n\right)\right)\right] && ((5.10), (5.5)) \\
&= 0. && \square
\end{aligned}$$

It turns out the condition in equation (5.5) holds when the coefficients have finite support.

Corollary 5.3.1. *Let $A_0, A_1, \ldots$ be a sequence of i.i.d. complex random variables with finite support that satisfy $\mathbf{P}(A_i = 0) = 0$. Let R_n be a random root of the polynomial*

$$p_n(z) = A_0 + A_1 z + \cdots + A_{n-1} z^{n-1} + z^n.$$

Then, as $n \to \infty$, the sequence of random variables $\operatorname{Arg}(R_n)$ converges in distribution to the uniform distribution on $[0, 2\pi)$.

Proof. Since the A_i are finitely supported, there is some bound B such that $\mathbf{P}(|A| < B) = 1$. By Lemma 4.2, for $w \geq 1$, the random variable $\mathbf{P}(|R_n^w| < 1+B^w) = 1$, so $|R_n^w|$ is also finitely supported. By Theorem 5.2, $|R_n^w| \xrightarrow{p} 1$, but since $|R_n^w|$ is finitely supported but this means $\mathbf{E}[|R_n^w|] \to 1$. The corollary then follows from Theorem 5.3. □

6 Conjectures from numerical simulations

6.1 Approximately uniform distribution of roots for a randomly-selected polynomial

One property we thought that the random polynomials might have is a uniform spread of roots, i.e. the roots of a randomly-chosen polynomial should not exhibit significant clustering with high probability as $n \to \infty$. Since in the limit $n \to \infty$ the magnitude of the roots approaches 1, the arguments of the roots are the important factor when considering their closeness. Therefore, we will formulate clustering of the roots in the limit in terms of the arguments of the roots.

To make this conjecture precise, for an n-th degree polynomial p and closed interval $I = [a, b] \subset [0, 2\pi]$, denote the number of roots of p whose argument is in I by $\#(p, I)$. We can consider the the quantity $c(p, I) = \frac{\#(p,I)/(b-a)}{n/2\pi}$ as a measure of the "concentration" of the arguments of the roots of p inside the interval I; it measures how big the actual number of roots whose arguments are in I is compared to the number that would be expected if the roots were distributed in a perfectly uniform manner. We can now formulate our conjecture.

Conjecture 6.1. *If $P_{(n)}(z) = A_0^{(n)} + \cdots + A_{n-1}^{(n)} z^{n-1} + z^n$ is a random n-th degree polynomial where the $A_i^{(n)}$ are i.i.d Gaussian distributions with mean zero, for any interval $I \subset [0, 2\pi]$ and any $\epsilon > 0$ we have that*

$$\lim_{n\to\infty} \mathbf{P}\left(|c(P^{(n)}, I) - 1| > \epsilon\right) = 0. \tag{6.1}$$

This can of course be conjectured for distributions other than Gaussian with mean zero. However, numerical evidence was only computed for the case of Gaussian. For a fixed interval width, random polynomials of increasing degree were generated, and the arguments of their roots were computed. Then, the *maximum* concentration in any interval of the given length was computed. As n increased, this maximum concentration decreased, seeming to approach 1. The results of this simulation for polynomials with complex Gaussian coefficients with variance 0.01 and interval length 0.1 are shown in Figure 6.1.

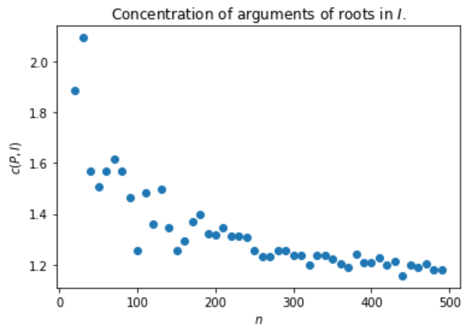

Figure 6.1: Maximum value of $c(P, I)$ v.s. n

6.2 Repulsion of roots

Another conjectured property related to the distribution of roots of a given randomly-selected polynomial is that the roots *repel* each other, that is the roots should tend to "spread out" from each other. Note that this is not the same as saying that there tends to be no clustering (for example roots of a polynomial p might be such that $c(p, I)$ is close to 1, but there are a few roots that are still very close together, or even equal). The way we measure repulsion is in terms of the variance of the number of roots whose arguments fall within a given subinterval I of $[0, 2\pi]$.

The intuition for why this captures the idea of repulsion is that it measures a sort of correlation between the roots. For example, suppose we are considering the distribution of roots of a random n-th degree polynomial p for large n so that the magnitudes roots are close to 1, and we are only interested in the distribution of arguments. Consider the extreme case that the distribution of roots is such that all n roots are identical; the roots would be in I with probability $\text{length}(I)/2\pi$ and outside with probability $1 - \text{length}(I)/2\pi$. The variance of the number of roots in I is then $\Theta(n^2)$. At the other extreme, if the distribution of roots were such that the roots were always perfectly uniformly spaced in their arguments, then for an interval there could only be either $\lfloor n * \text{length}(I)/2\pi \rfloor$ or $\lfloor n * \text{length}(I)/2\pi \rfloor + 1$. The variance in this case is $\Theta(1)$. Finally, if it were the case that the roots were independently uniformly distributed, then the number of roots in I would be a binomial random variable with parameter $\text{length}(I)/2\pi$ which has $\Theta(n)$ variance. Hence, a variance of $\Theta(n^2)$ would correspond to an maximum "attraction" of roots, variance $\Theta(n)$ to an "independence" of roots, and variance $\Theta(1)$ to a maximum "repulsion" of roots.

Now, if this variance were computed for a specific distribution of random coefficients, its asymptotic behavior would give a measure of the amount of repulsion or attraction between the roots of a randomly generated polynomial.

Numerical simulation suggests that the growth of the variance may be $\Theta(\log n)$ or $\Theta(n^\alpha)$ for $\alpha < 1$, which can reasonably be considered to be much closer to $\Theta(1)$ than to $\Theta(n)$. The graphs for the estimated variance for increasing n are shown in Figure 6.2 for both complex Gaussian coefficients with variance 0.01 and coefficients which are the uniform distribution on the two values $\{+1, -1\}$. The variances were estimated by computing the roots of 50 random polynomials for select degrees, then finding the variance of the resulting set.

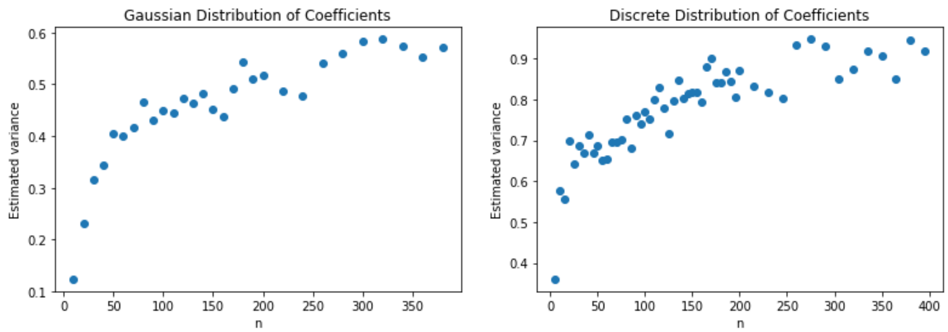

Figure 6.2: Estimated variance of $\#(P^{(n)}, [0, 0.1])$ v.s. n

6.3 Distribution of the magnitude of max root approaches fixed distribution as n increases

As we were exploring the distribution of the roots of the random polynomial, we also explored the distribution of just the maximum magnitude root. For these simulations, we sampled 5000 polynomials of degrees $n = 1$, 3, and 10. The coefficients A_i were sampled from complex Gaussians with mean 0 and standard deviation 1. For each sampled polynomial, we numerically computed its roots, and plotted a histogram of the root with the maximum magnitude, over all 5000 sampled polynomials, for each of the three n.

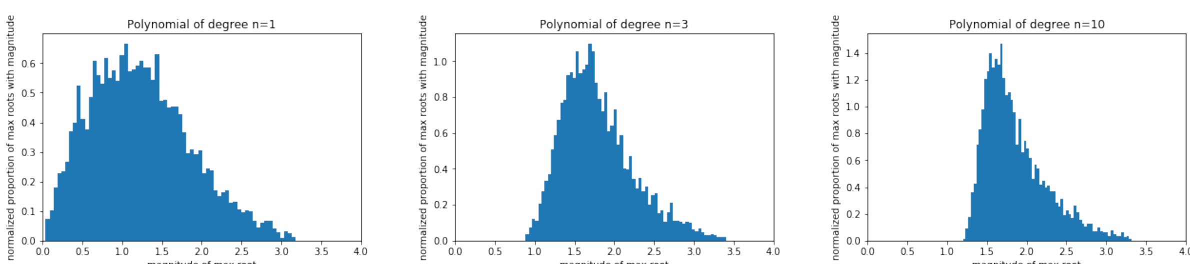

Figure 6.3: Distribution of the magnitude of the root with maximum magnitude over complex normal perturbed polynomials of degrees 1, 3, and 10, respectively.

Figure 6.3 suggests that the distribution of these maximum root magnitudes converge to some fixed distribution as the degree n further increases to infinity. We did not run any simulations for different coefficient distributions A_i; however, we suspect that the A_i must at least remain i.i.d. for this conjecture to hold.

References

[1] Michael Artin. *Algebra, 2nd Edition*. Pearson Prentice Hall, 2011.

[2] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.