Separation of roots of random polynomials

Marcus Michelen

joint with Oren Yakir (MIT)

Northwestern University
Department of Mathematics

Oxford Discrete Mathematics and Probability Seminar
November 2025

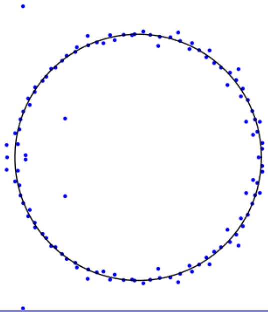
Random polynomials

Let $f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ be a degree n polynomial whose coefficients ξ_j are independent and identically distributed random variables taken uniformly from $\{-1, 1\}$. (or $\xi_j \sim \mathcal{N}(0,1)$ or just $\mathbb{E}\xi_j = 0$ and $\mathrm{Var}(\xi_j) = 1$)

The polynomial f_n has n roots in $\mathbb{C}$ (counting by multiplicity).

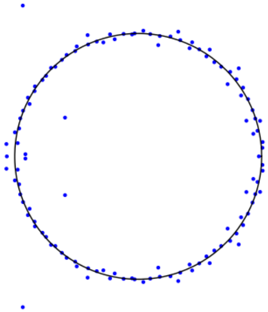
Guiding Question

What can we say about the roots when n is large?



Unpacking the picture

$f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ where $\xi_j \in \{-1, 1\}$ uniformly at random.



1. Most roots (but not all!) are close to the unit circle
2. Since the coefficients are real, there is symmetry about the real axis/x-axis. Some roots are real (there are 2) real roots here.
3. The roots appear to repel each other.

Our plan

Here's how the rest of the talk will go:

1. (Mostly classical) theorems that rigorously describe the behavior from this picture.
2. Isolating a single tool: Erdős's solution to the Littlewood-Offord problem, and outline two proofs, one using extremal combinatorics and one using fourier analysis.
3. What did we prove and how?
4. Open problems, future directions

The first questions answered: the number of real roots

Many classical predecessors Bloch-Polya, Littlewood-Offord, first sharp result is: Kac 1943 proved when the coefficients are $\mathcal{N}(0,1)$ the number of real roots is $(\frac{2}{\pi} + o(1))\log n$. “In case the [coefficients] are not normally distributed (but all have the same distribution with standard deviation 1) one can still prove [this result].”

In 1946 Kac extended to other continuous distributions like uniform in $[-1,1]$ and stated “Upon a closer examination it turns out that the proof which I had in mind, based on the central limit theorem of the calculus of probability, is inapplicable to the discrete case.”

Erdős-Offord (1956) proved it for $\{-1,1\}$ coefficients (and many other variables).

This is a **universality** phenomenon: the behavior for the number of real roots doesn't depend much on the actual distribution of coefficients.

The bulk behavior: most are near the unit circle

Let $\mathcal{Z}$ denote the zero (multi)set of f_n. Then

$$\frac{|\{\alpha \in \mathcal{Z} : |\alpha| \in [1-\varepsilon, 1+\varepsilon]\}|}{n} \to 1$$

$$\frac{|\{\alpha \in \mathcal{Z} : \arg(\alpha) \in [a,b]\}|}{n} \to \frac{b-a}{2\pi}$$

Theorem (Erdős-Turán, 1950)

This is true deterministically for polynomials whose coefficients are on the same exponential scale (i.e. $|\xi_j|^2/|\xi_0\xi_n| = e^{o(n)}$ for all j).

Same as saying the probability measure $\mu_n = \frac{1}{n}\sum_{\alpha\in\mathcal{Z}}\delta_\alpha$ converges to the uniform distribution on the unit circle. To prove this, it is sufficient to prove $\frac{1}{n}\log|f_n(z)|$ converges to 0 for all $|z| < 1$ and to $\log|z|$ for all $|z| > 1$. This is since $\mu_n = \frac{1}{2\pi}\Delta \log|f_n|$.

Most roots are $\Theta(1/n)$ away from the unit circle (Shepp-Vanderbei '95, Konyagin-Schlag '99, Ibragimov-Zeitouni, '97)

Roots indeed to repel each other

Most roots are within $1/n$ away from the unit circle and there are n roots. So if $t > 0$ and $z = 1 + O(1/n)$ then one expects

$$\mathbb{P}(\exists\ \alpha \in \mathcal{Z} : |z - \alpha| \leq t/n) = \Theta_t(1)$$

which turns out to be true. Set $B_t(z) = \{w : |z - w| \leq t/n\}$. Then in fact

$$\mathbb{P}(\text{at least two roots in } B_t(z)) \ll \mathbb{P}(\text{at least one root in } B_t(z))^2$$

for $t \in [n^{-c}, o(1)]$.

Credit to various works of Shiffman-Zelditch for Gaussians and universality works of Tao-Vu.

A single tool: the Littlewood-Offord problem

Relevant for all pieces is understanding the probability $f(z)$ is small. Let's fix z with $|z| \approx 1$ and look at $\mathbb{P}(|f(z)| \leq 1)$. Note that $\mathbb{E}|f(z)|^2 \approx n$ so typically $|f(z)| = \Theta(n^{1/2})$.

We have

$$f(z) = \xi_n z^n + \xi_{n-1} z^{n-1} + \cdots + \xi_1 z + \xi_0 \,.$$

More generally let's look at $X = a_n\xi_n + a_{n-1}\xi_{n-1} + \cdots + a_1\xi_1$ where $\xi_j \in \{-1, 1\}$ uniformly at random and $|a_n| \geq 1$. How big can $\mathbb{P}(|X| < 1)$ be? For $a_j \equiv 1$ we have $\mathbb{P}(|X| < 1) = 2^{-n}\binom{n}{n/2} \asymp n^{-1/2}$.

Theorem (Erdős's solution to the Littlewood-Offord problem, 1945)

For all $|a_j| \geq 1$ we have $\mathbb{P}(|X| < 1) \leq \frac{C}{\sqrt{n}}$.

We will see two proofs: one using extremal combinatorics (Erdős's original proof) and a fourier analytic proof.

Erdős and Littlewood-Offord via extremal combinatorics

Theorem (Erdős)

Let $a_j \in \mathbb{R}$ with $|a_j| \geq 1$ and set $X = \sum_{j=1}^n a_j \xi_j$ where ξ_j are i.i.d. and uniformly distributed in $\{-1, 1\}$. Then $\mathbb{P}(|X| < 1) \leq 2^{-n}\binom{n}{\lfloor n/2 \rfloor} \leq \frac{C}{n^{1/2}}$.

Assume without loss of generality that $a_j \geq 1$.

Associate to an instance $\{\xi_j\} \in \{-1, 1\}^n$ the set $S \subset [n]$ of indices giving $+1$; so $S = \{i : \xi_i = +1\}$. We can think of $X = X(S)$.

Note that if $T \subsetneq S$ then $X(S) - X(T) \geq 2$. In particular: if $|X(S)| < 1$ and $T \subsetneq S$ (or $S \subsetneq T$) then $|X(T)| > 1$.

This means that $\{S \subset [n] : |X(S)| < 1\}$ is an *antichain* in the hypercube, so Sperner's lemma states $|\{S \subset [n] : |X(S)| < 1\}| \leq \binom{n}{\lfloor n/2 \rfloor}$. $\square$

Erdős and Littlewood-Offord via fourier analysis

Theorem

Let $a_j \in \mathbb{R}$ with $|a_j| \in [1, 10]$ and set $X = \sum_{j=1}^n a_j \xi_j$ where ξ_j are i.i.d. and uniformly distributed in $\{-1, 1\}$ (or more generally just non-constant). Then $\mathbb{P}(|X| < 1) \leq \frac{C}{n^{1/2}}$.

Set $g(x) = \mathbf{1}\{x \in (-1,1)\}$ so $\mathbb{P}(|X| < 1) = \mathbb{E}_X g(X)$.

$$g \overset{\text{F.T.}}{\longleftrightarrow} \frac{\sin(\theta)}{\theta}, \qquad (g * g) \overset{\text{F.T.}}{\longleftrightarrow} \left(\frac{\sin(\theta)}{\theta}\right)^2$$

and note that $g * g$ has support $[-2, 2]$. We also can bound $g(x) \leq 2(\sin(x)/x)^2$.

$$\mathbb{E}_X g(X) \lesssim \mathbb{E}_X \left(\tfrac{\sin X}{X}\right)^2 = c \int (g * g)(\theta)\, \mathbb{E}_X e^{i\theta X}\, d\theta \lesssim \int_{-2}^{2} |\mathbb{E}_X e^{i\theta X}|\, d\theta$$

$$= \int_{-2}^{2} \prod_j |\cos(a_j \theta)|\, d\theta \leq \int_{-2}^{2} \exp(-cn\theta^2)\, d\theta \lesssim n^{-1/2}. \qquad \square$$

Returning to our random polynomial: separation of roots

So: $f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ where $\xi_j \in \{-1, 1\}$ are independent and chosen uniformly at random, where $\mathcal{Z} = \{\alpha_j\}_{j=1}^{n}$ is the (multi)set of zeros.

Recall *most* roots are near the unit circle (in fact $\Theta(1/n)$) away from the unit circle.

The roots experience repulsion. How can we quantify that?

Set $m_n = \min_{i<j} |\alpha_i - \alpha_j|$ to be the minimal separation of roots of f_n.

Theorem (M.-Yakir, 2025)

We have $m_n = \Theta(n^{-5/4})$ with high probability. In fact $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$ where $c_* > 0$ and does not depend on the coefficient distribution.

Two quantitative bits here: the power of $5/4$ on n and the power of 4 on λ. Both capture repulsion between the roots....let's see why.

A toy model: what if there were no repulsion?

$f_n(z) = \sum_{j=0}^n \xi_j z^j$, with roots α_j set $m_n = \min |\alpha_i - \alpha_j|$. Then $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$

Why does this capture repulsion? Let's consider a toy model where I place n independent points $X_1, \ldots, X_n$ in $\mathcal{A} = \{z : 1 - 1/n \leq |z| \leq 1 + 1/n\}$ independently and uniformly at random. How does the separation $M_n = \min |X_i - X_j|$ behave?

Let $\varepsilon > 0$ be small, let's compute the expected number of pairs with $|X_i - X_j| \leq \varepsilon$.

For a given $z \in \mathcal{A}$ the expected number of pairs in $B_\varepsilon(z)$ is

$$\mathbb{E}[(i,j) : X_i, X_j \in B_\varepsilon(z)] \approx n^2 \mathbb{P}(X_1, X_2 \in B_\varepsilon(z)) = n^2 \mathbb{P}(X_1 \in B_\varepsilon(z))^2 \approx n^4 \varepsilon^4.$$

It takes $\approx n^{-1}\varepsilon^{-2}$ balls to cover $\mathcal{A}$ so the expected number of pairs with distance $\leq \varepsilon$ is $n^3 \varepsilon^2$.

If $n^3 \varepsilon^2 = \Theta(1)$ then expect the number of pairs at distance ε to be Poisson of mean $\Theta(n^3 \varepsilon^2)$. $\mathbb{P}(M_n \geq \lambda n^{-3/2}) \to \exp(-c' \lambda^2)$.

How is the actual calculation different?

$f_n(z) = \sum_{j=0}^n \xi_j z^j$, with roots α_j set $m_n = \min|\alpha_i - \alpha_j|$. Then $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$

For independent points, we saw $\mathbb{E}[(i,j) : X_i, X_j \in B_\varepsilon(z)] \approx n^4 \varepsilon^4$.

Theorem (M.-Yakir, 2025)

For $z = 1 + \Theta(1/n)$ we have $\mathbb{E}[(i,j) : \alpha_i, \alpha_j \in B_\varepsilon(z)] \approx n^5 \varepsilon^6$.

....but for what $\varepsilon > 0$? As we saw, this turns out to be true for $\varepsilon = [n^{-c-1}, o(n^{-1})]$...but we need better than that. A quick calculation is that $\mathbb{P}(f_n(1) = f_n'(1) = 0) = \Theta(n^{-2})$ so we can't just take all $\varepsilon > 0$...at least not for all z.

It turns out that the problem with $z = 1$ is that $\arg(z) = 0$ is close to a rational number of small denominator. We could take ε smaller for $z = i$ and smaller for $z = e^{i\pi/4}$. If we want the above to hold for $\varepsilon \approx n^{-A}$ it turns out we can just omit z that is close to a $O_A(1)$ root of unity.

An approximate outline of the whole argument

$f_n(z) = \sum_{j=0}^{n} \xi_j z^j$, with roots α_j set $m_n = \min |\alpha_i - \alpha_j|$. Then $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$.

A sketch to prove that if $Y_\lambda = \{i < j : |\alpha_i - \alpha_j| \leq \lambda n^{-5/4}\}$ then Y_λ converges to a Poisson random variable of mean $\mu = c_* \lambda^4$.

- We will use the method of moments, so we want to show that for each k we have $\mathbb{E}[Y_\lambda(Y_\lambda - 1) \cdots (Y_\lambda - (k-1))] \to \mu^k$.
- Omit pairs with $||z| - 1| \geq \widetilde{\Omega}(1/n)$. [more on this, time permitting]
- Omit pairs where z is close to an $O(1)$-root of unity.
- For points $z_1, \ldots, z_k = 1 + O(1/n)$ separated by $\gg 1/n$ we have that the probability there are pairs with distance $\lambda n^{-5/4}$ near each of the k points approximately factors. This is done by comparing to a Gaussian, like a beefed up version of the Fourier proof of Littlewood-Offord. [more on this, next slide]

Detecting roots and detecting pairs

Given a point $z = 1 + O(1/n)$, how can I tell if there is a root near z? Taylor expand near z:

$$f(w) \approx f(z) + (w - z)f'(z) + \mathrm{Error}\,.$$

If $f(z)/f'(z)$ is atypically small and f'' is typical, then we expect a zero of f at $w \approx z - f(z)/f'(z)$. To see when there are *two* roots near z, need some event depending on $(f(z), f'(z), f''(z))$.

To calculate probabilities related to $(f(z), f'(z), f''(z))$ we prove a local central limit theorem / Gaussian comparison principle that gets stronger when z is far from roots of unity.

But we actually see something extra here: if $f'(z)$ is large, there is *more likely* to be a root near z....this tells us where repulsion comes from!

Why do roots repel each other

We saw there is a root near z if $f(z)/f'(z)$ is small. So if f' is large, there is *more likely* to be a root. So if I take a typical root α of f, it is *unlikely* for $f'(\alpha)$ to be small.

A thought experiment: you are watching 10 runners on a circular track and 5 are very fast and 5 are slow; you see the fast runners more often.

Why does this imply roots repel? If two roots α, β are close, then Rolle's theorem / Mean value theorem says that f' is zero somewhere in between...but if $|\alpha - \beta|$ is small, then this means $f'(\alpha)$ is small. There is a tension:

[Two close roots $\implies$ f' is small on a root] vs. [f' is typically not small on roots]

So roots typically repel.Same story holds for many random functions, e.g. eigenvalues of random matrices.

One last wrinkle: eliminating roots far from the unit circle

The behavior of f is *not* universal for $|z|$ far from 1. Let's focus on $|z| < 1$ and consider the $n = \infty$ limit: $f_\infty(z) = \sum_{j \geq 0} \xi_j z^j$ which is analytic for $|z| < 1$. The random variable $f_\infty(z)$ is *not* like a gaussian...for instance $f_\infty(\frac{2}{1+\sqrt{5}})$ is not even a continuous random variable! (Erdős, 1939)

We need a separate approach.

Theorem (M.-Yakir, 2025)

Let ξ_j be independent and identically distributed random variables with $\mathbb{E}\log(1 + |\xi_j|) < \infty$, then with probability 1 the power series f_∞ is analytic in $|z| < 1$ and has no double roots except perhaps at the origin.

Proof is perturbative: given the first M coefficients, it is unlikely that the rest will combine to give you a double root. Explicitly uses the solution to the Littlewood-Offord problem.

Outline revisited

$f_n(z) = \sum_{j=0}^{n} \xi_j z^j$, with roots α_j set $m_n = \min |\alpha_i - \alpha_j|$. Then $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$.

A sketch to prove that if $Y_\lambda = \{i < j : |\alpha_i - \alpha_j| \leq \lambda n^{-5/4}\}$ then Y_λ converges to a Poisson random variable of mean $\mu = c_* \lambda^4$.

- We will use the method of moments, so we want to show that for each k we have $\mathbb{E}[Y_\lambda (Y_\lambda - 1) \cdots (Y_\lambda - (k-1))] \to \mu^k$.
- Omit pairs with $||z| - 1| \geq \widetilde{\Omega}(1/n)$ by a perturbative argument on the infinite power series corresponding to $n = \infty$.
- Omit pairs where z is close to an $O(1)$-root of unity via Littlewood-Offord approach.
- For points $z_1, \ldots, z_k = 1 + O(1/n)$ separated by $\gg 1/n$ we have that the probability there are pairs with distance $\lambda n^{-5/4}$ near each of the k points approximately factors. This is done by taylor expanding to second order and comparing $((f(z_j), f'(z_j), f''(z_j))_{j=1}^{k}$ to a Gaussian vector (using that none is close to a $O(1)$-root of unity).

Open problems and future directions: repeated roots

Our theorem holds provided the variables are *sub-gaussian* meaning $\mathbb{P}(|\xi_j| \geq t) \leq e^{-ct^2}$ for some $c > 0$. We also have the following corollary of our theorem:

Corollary

Let ξ_j be independent and subgaussian and set $f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ then $\mathbb{P}(f_n$ has a double root other than at $0) = o(1)$.

This was known for $\xi \in \{-1, 1\}$ and some other integer-valued distributions by Peled-Sen-Zeitouni, Feldheim-Sen but is new in this generality.

Conjecture

Let ξ be a non-constant random variable and set $f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ then $\mathbb{P}(f_n$ has a double root other than at $0) = o(1)$.

Much harder to do something as analytic in this generality, perhaps a perturbative approach and a small bit of algebra might be useful.

Open problems and future directions: discriminant

Part of our motivation was for studying the discriminant: if $f_n(z) = \sum_{j=0}^n \xi_j z^j$ then $|\mathrm{disc}(f_n)| = |\xi_n|^{2n-2} \prod_{i<j} |\alpha_i - \alpha_j|^2$.

Theorem (M.-Yakir, 2025)

Let ξ_j be independent, mean 0, variance 1 and subgaussian. Then

$$\frac{\log |\mathrm{disc}(f_n)| - 2n \log n}{n} \xrightarrow[\mathbb{P}]{n\to\infty} D_*$$

for some universal $D_* \neq 0$. So $|\mathrm{disc}(f_n)| = n^{2n} e^{(D_* + o(1))n}$.

Conjecture (Bary-Soroker + Kozma)

Let $\xi_j \in \{0,1\}$ uniformly at random (or $\{-1,1\}$ uniformly at random etc.). Then with high probability $|\mathrm{disc}(f_n)|$ is not a perfect square.

[$|\mathrm{disc}(f_n)|$ is a perfect square if and only if $\mathrm{Gal}(f_n) \leq A_n$]

Summary

We consider $f_n(z) = \sum_{j=0}^{n} \xi_j z^j$ with $\xi_j \in \{-1, 1\}$ uniformly at random (for instance).

Classical work of Erdős-Turán (and others) show most roots are near the unit circle.

The roots experience repulsion.

Theorem (M.-Yakir, 2025)

Set $m_n = \min_{i<j} |\alpha_i - \alpha_j|$ *where* $\{\alpha_j\}_{j=1}^{n}$ *are the roots of* f_n*. Then* $m_n = \Theta(n^{-5/4})$ *and in particular* $\mathbb{P}(m_n \geq \lambda n^{-5/4}) \to \exp(-c_* \lambda^4)$*.*

Thank you!