Electron. Commun. Probab. **29** (2024), article no. 27, 1–10.
https://doi.org/10.1214/24-ECP596
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Almost sure behavior of the zeros of iterated derivatives of random polynomials

Marcus Michelen* Xuan-Truong Vu†

Abstract

Let $Z_1, Z_2, \ldots$ be independent and identically distributed complex random variables with common distribution μ and set

$$P_n(z) := (z - Z_1)\cdots(z - Z_n)\,.$$

Recently, Angst, Malicet and Poly proved that the critical points of P_n converge in an almost-sure sense to the measure μ as n tends to infinity, thereby confirming a conjecture of Cheung-Ng-Yam and Kabluchko. In this short note, we prove for any fixed $k \in \mathbb{N}$, the empirical measure of zeros of the kth derivative of P_n converges to μ in the almost sure sense, as conjectured by Angst-Malicet-Poly.

Keywords: random polynomial; iterated derivatives; anti-concentration.
MSC2020 subject classifications: 60B99; 30C15.
Submitted to ECP on August 30, 2023, final version accepted on May 13, 2024.

1 Introduction

Let μ be a probability measure on $\mathbb{C}$ and let $(Z_j)_{j\geqslant 1}$ be independent and identically distributed (i.i.d.) complex random variables with distribution μ. Define the sequence of random polynomials $(P_n)_{n\geqslant 1}$ via

$$P_n(z) := (z - Z_1)\cdots(z - Z_n)\,. \tag{1.1}$$

Pemantle and Rivin [19] introduced this model and conjectured that the critical points of P_n are close to the roots of P_n. More rigorously, for each fixed $k \in \mathbb{N}$ let $\nu_n^{(k)}$ to be the empirical measure of $P_n^{(k)}$, i.e.

$$\nu_n^{(k)} := \frac{1}{n-k} \sum_{z\in\mathbb{C}:P_n^{(k)}(z)=0} \delta_z\,, \tag{1.2}$$

and let μ_n denote the *empirical measure* of P_n, i.e.

$$\mu_n := \frac{1}{n} \sum_{z\in\mathbb{C}:P_n(z)=0} \delta_z\,, \tag{1.3}$$

where δ_y denotes the point mass at y.

*University of Illinois Chicago, United States of America. E-mail: michelen@uic.edu
†University of Illinois Chicago, United States of America. E-mail: tvu25@uic.edu

Pemantle and Rivin conjectured that in the case of $k = 1$, we have $\nu_n^{(1)}$ converges weakly to μ; Pemantle and Rivin proved this under the assumption that μ has finite 1-energy. Kabluchko [10] proved Pemantle and Rivin's conjecture, showing that $\nu_n^{(1)} \to \mu$ in probability as $n \to \infty$. Kabluchko's result was extended by Byun, Lee and Reddy [3] who proved that for each fixed $k \in \mathbb{N}$ one has that $\nu_n^{(k)}$ converges weakly to μ in probability. Recently, the authors showed the same holds if k grows slightly slower than logarithmically in n [16]. The works [3] and [16] on convergence of higher derivatives follow the same general strategy as Kabluchko's original proof [10], which much of the new ingredients coming in to handle an anti-concentration estimate. For more references on this model and adjacent models, see the works [1, 4, 5, 8, 9, 11, 12, 18, 20, 21] and the references therein.

Cheung-Ng-Yam [4] and independently Kabluchko (see [16]) conjectured that in fact $\nu_n^{(1)}$ should weakly converge to μ *almost surely* and not just in probability. This was recently proven by Angst, Malicet and Poly [2] for all probability measures μ. Angst, Malicet and Poly also conjectured [2] that the almost-sure convergence of $\nu_n^{(k)}$ should hold for each fixed $k \in \mathbb{N}$. In this short note, we confirm their conjecture.

Theorem 1.1. *Almost surely with respect to $\mathbb{P}$, for each fixed $k \in \mathbb{N}$ the sequence of empirical measure $\nu_n^{(k)}$ converges to μ as n tends to infinity.*

Our proof of Theorem 1.1 takes inspiration from Angst, Malicet and Poly's proof of the $k = 1$ case [2]; the new ingredient is to handle a non-linear, high-dimensional, multivariate anti-concentration problem via a decoupling approach. We first outline the general shape of the Angst-Malicet-Poly approach as well as where our contribution comes into play in Section 1.1. We then prove our anti-concentration estimate in Section 2 and complete the proof of Theorem 1.1 in Section 3.

1.1 Outline of the Angst-Malicet-Poly strategy and our contribution

The main engine behind Angst-Malicet-Poly is a simple-yet-powerful fact about probability measures on the Riemann sphere. To set up their Lemma, set $\widehat{\mathbb{C}}$ to be the Riemann sphere, let $\mathcal{M} := \{\psi(z) = \frac{az+b}{cz+d}\}$ be the set of Möbius transformations, and let $\lambda_{\mathcal{M}}$ be the measure on $\mathcal{M}$ inherited by setting taking the complex Lebesgue measure on the tuple (a, b, c, d). Define $\log_- z = |\log z| \mathbf{1}_{z\in[0,1]}$. Their main engine is the following lemma [2, Lemma 2.7]:

Lemma 1.2. *Let $\widehat{m}_1$ and $\widehat{m}_2$ be two probability measures on $\widehat{\mathbb{C}}$ so that*

$$\int_{\widehat{\mathbb{C}}} \log_- |\psi(z)| \, d\widehat{m}_1(z) \leqslant \int_{\widehat{\mathbb{C}}} \log_- |\psi(z)| \, d\widehat{m}_2(z) \tag{1.4}$$

for almost-every Möbius transformation $\psi \in \mathcal{M}$. Then $\widehat{m}_1 = \widehat{m}_2$.

An appealing aspect of this Lemma is that it requires only a one-sided bound; further, the space of probability measures on $\widehat{\mathbb{C}}$ is compact, and so it will suffice to verify (1.4) where $\widehat{m}_1$ will be an arbitrary cluster point $\widehat{\nu}_\infty$ of the sequence $(\nu_n^{(k)})_{n\geqslant 1}$ and $\widehat{m}_2$ will be the measure μ.

The route towards establishing (1.4) begins at Jensen's formula. Letting $C(0, 1)$ denote the unit circle, one may apply Jensen's formula to the ratio $S_n := \frac{P_n^{(k)}}{k! P_n}$ to obtain

$$\sum_\rho \log_- |\psi(\rho)| - \sum_\zeta \log_- |\psi(\zeta)| \leqslant \max_{z\in\psi^{-1}(C(0,1))} \log |S_n(z)| - \log |S_n(\psi^{-1}(0))| \tag{1.5}$$

where $\{\rho\}$ enumerates the roots of $P^{(k)}$ and $\{\zeta\}$ enumerates the roots of P (see Fact 3.1). Our task then is to control the right-hand side. The term with the maximum is fairly straightforward to control almost-surely (Lemma 3.2), and it is the term $\log |S_n(\psi^{-1}(0)|$ that is more challenging. In particular, if we set $a = \psi^{-1}(0)$ then

$$S_n(a) = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \frac{1}{a - Z_{i_1}} \cdots \frac{1}{a - Z_{i_k}} \,. \tag{1.6}$$

The case of $k = 1$, this is precisely a sum of i.i.d. random variables and so we depending on the distribution of Z and the choice of a we may have that $\mathbb{P}(S_n = 0) = \Theta(n^{-1/2})$. Since we seek almost sure statements, this is too large to apply Borel-Cantelli. To get around this issue, Angst-Malicet-Poly look instead at *triples* of Möbius transformations. For most such triples (ψ_1, ψ_2, ψ_3) one has that the vector $(\psi_1^{-1}(0), \psi_2^{-1}(0), \psi_3^{-1}(0))$ consists of three distinct complex numbers, say (a, b, c). The vector $(S_n(a), S_n(b), S_n(c))$ now behaves like a sum of three-dimensional random variables. In particular, a sufficiently general version of Erdős's solution to the Littlewood-Offord problem shows that the probability all coordinates of $(S_n(a), S_n(b), S_n(c))$ are small simultaneously decays like $O(n^{-3/2})$, which is now summable. An application of Borel-Cantelli will allow one to deduce that almost-surely for generic triples of Möbius transformations and all large enough n we have at least one of $(S_n(\psi_1^{-1}(0)), S_n(\psi_2^{-1}(0)), S_n(\psi_3^{-1}(0)))$ is at least, say, 1 in modulus. Working with a given cluster point $\widehat{\nu}_\infty$ of $(\nu_n^{(k)})_{n\geqslant 1}$ and applying (1.5) together with an application of the law of large numbers to handle the sum over $\{\zeta\}$ will prove (1.4).

The main challenge in adapting this approach to fixed $k \in \mathbb{N}$ is to handle the anti-concentration estimate. In particular, for fixed $k \geqslant 2$, handling the quantity $\mathbb{P}(S_n(a) = 0)$ is a non-linear anti-concentration problem, and major open problems remain in this arena. As an example, one expects that for each $k \geqslant 2$ one has $\mathbb{P}(S_n(a) = 0) = O(n^{-1/2})$, but this is only known up to subpolynomial factors [15] for $k \geqslant 3$; the case of $k = 2$ was recently solved by a work of Kwan-Sauermann. Furthermore, we need to consider vectors of such quantities. Roughly, for each fixed k, we need to take L large enough so that for distinct complex numbers $(z_1, \ldots, z_L)$ we have

$$\sum_{n \geqslant 1} \mathbb{P}(|S_n(z_1)| \leqslant 1, \ldots, |S_n(z_L)| \leqslant 1) < \infty \,.$$

To handle this quantity, we use a *decoupling* approach for anti-concentration. This was introduced by Costello-Tao-Vu [7] in their study of random symmetric matrices and anti-concentration of quadratic forms (see also the survey [17]). The intuition here is to tackle multilinear anti-concentration problems by comparing them to linear anti-concentration problems, at the cost of decreasing the rate of decay of the resulting bounds. For us, we need to apply a decoupling lemma to the vector $(S_n(z_1), \ldots, S_n(z_L))$ and handle all coordinates simultaneously in order to obtain a high-dimensional but *linear* anti-concentration problem. Our main new contribution is the following Lemma:

Lemma 1.3. *Suppose μ does not have finite support and let $k \in \mathbb{N}$. Then for $L = 2^{k+2}k$ and all pairwise distinct complex numbers $(z_1, \ldots, z_L)$ we have*

$$\mathbb{P}(|S_n(z_j)| \leqslant 1 \text{ for all } j \in [L]) \leqslant \frac{C}{n^2}$$

where $C > 0$ depends on k and μ.

We note that increasing L yields an increase in the exponent on n on the right-hand side; we only need the right-hand side to be summable in n. Further, it is plausible that the right-hand side of Lemma 1.3 should be of the order $n^{-L/2}$; however, even in the case of $k = 2$ and $L = 1$ this is a non-trivial instance of a significant open problem known as the quadratic Littlewood-Offord problem (see [6, 7, 13, 15]). Since we only need summability, the sub-optimal bounds attained by decoupling will be strong enough provided we take L large as in Lemma 1.3.

We note that this approach differs fundamentally from the anti-concentration approach of our previous work [16]. In [16] we deduced our anti-concentration estimate from a powerful theorem of Meka-Nguyen-Vu [15] which in turn is proven by a sophisticated Gaussian comparison argument.

The decoupling approach and proof of Lemma 1.3 is handled in Section 2. We then prove Theorem 1.1 in Section 3, and import the necessary tools and adapt ideas from [2].

1.2 Notation

Throughout, the random variables $(Z_j)_{j\geqslant 1}$ are defined on the common probability space $(\mathbb{P},\mathcal{F},\Omega)$. The random polynomials we consider are defined by $P_n(z) = (z - Z_1)\cdots(z-Z_n)$. The measure μ_n is the empirical measure of P_n and the measure $\nu_n^{(k)}$ is the empirical measure of $P_n^{(k)}$. We will make use of the ratio $S_n = \frac{P_n^{(k)}}{k!P_n}$.

We set $\lambda_{\mathbb{R}}$ to be the Lebesgue measure on $\mathbb{R}$ and $\lambda_{\mathbb{C}}$ to be the Lebesgue measure on $\mathbb{C}$; we write $\widehat{\mathbb{C}}$ for the Riemann sphere. We denote $\mathcal{M} = \{\psi(z) = \frac{az+b}{cz+d}\}$ for the set of Möbius transformations and endow $\mathcal{M}$ with the measure $\lambda_{\mathcal{M}}$ induced by the taking the Lebesgue measure $\lambda_{\mathbb{C}}^{\otimes 4}$ on the tuples (a,b,c,d) defining the Möbius transformations. We denote $\log z = \log_+ z - \log_- z$ where

$$\log_- z = \begin{cases} |\log z|, & 0\leqslant z\leqslant 1,\\ 0, & z\geqslant 1,\end{cases} \quad \text{and} \quad \log_+ z = \begin{cases} 0, & 0\leqslant z\leqslant 1,\\ \log z, & z\geqslant 1,\end{cases}$$

where $\log_- 0 = +\infty$. We write $C(a,r)$ for the circle centered at $a\in\mathbb{C}$ of radius r. For $m\in\mathbb{N}$ we write $[m] = \{1,2,\dots,m\}$.

2 Anticoncentration via decoupling

The goal of this section is to prove Lemma 1.3; we begin with the abstract decoupling lemma of Costello, Tao and Vu [7].

Lemma 2.1. *Let $(Y_1,\dots,Y_r)$ be a collection of random variables taking values in an arbitrary measurable space and let $E = E(Y_1,\dots,Y_r)$ be an event depending on these variables. Set $(Y_1',\dots,Y_r')$ to be an independent copy of the collection of random variables $(Y_1,\dots,Y_r)$ with the same joint distribution. Then*

$$\mathbb{P}(E(Y_1,\dots,Y_r)) \leqslant \mathbb{P}\left(\bigwedge_{\alpha\subset[r]} E(Y_1^\alpha,\dots,Y_r^\alpha)\right)^{1/2^r}$$

where $Y_j^\alpha = Y_j$ if $j\in\alpha$ and $Y_j^\alpha = Y_j'$ if $j\notin\alpha$.

The strategy will be to apply this lemma and subsequently take linear combinations of various versions of $S_n(z_j)$ in order to obtain a linear inequality rather than a multi-linear inequality. We then will need a high-dimensional (linear) anti-concentration statement which is stated in [2]. A random vector $(X_1,\dots,X_d)\in\mathbb{C}^d$ is *non-degenerate* if there do not exist complex numbers α_j,β so that

$$\sum_{j=1}^d \alpha_j X_j - \beta = 0$$

almost surely. This non-degeneracy assumption asserts that $(X_1,\dots,X_d)$ is genuinely d-dimensional.

Proposition 2.2 (Proposition 2.1 from [2]). *Let $(X^n)_{n\geqslant 1} = (X_1^n, \ldots, X_d^n)_{n\geqslant 1}$ be a sequence of i.i.d. non-degenerate random vectors taking values in $\mathbb{C}^d$ and set $S_n = \sum_{k=1}^n X^k$. Then there is a constant C depending on d, r and the law of X so that for all n we have*

$$\sup_{x\in\mathbb{C}^d} \mathbb{P}\left(\|S_n - x\| \leqslant r\right) \leqslant \frac{C}{n^{d/2}}\,.$$

We are now ready to set up the decoupling approach. Recall that $(Z_j)_{j\geqslant 1}$ are the i.i.d. samples from μ giving the roots of the polynomials $(P_n)_{n\geqslant 1}$. For fixed n and k, partition $[n]$ into k disjoint sets $R_1, \ldots, R_k$ with $\lfloor n/k \rfloor \leqslant |R_j| \leqslant \lceil n/k \rceil$. For $j \in [k]$ define $Y_j = (Z_i)_{i\in R_j}$. We now think of the rational function $S_n(z)$ as a function not only of z but also of the quantities $(Y_j)_{j\in[k]}$ and so we write $S_n(z; Y_1, \ldots, Y_k)$ when we want to make this dependence explicit.

Applying Lemma 2.1 shows

$$\mathbb{P}(|S_n(z_j)| \leqslant 1 \text{ for all } j \in [L]) \leqslant \mathbb{P}\left(|S_n(z_j; Y_1^\alpha, \ldots, Y_k^\alpha)| \leqslant 1 \text{ for all } j \in [L], \alpha \subset [k]\right)^{1/2^k}\,. \tag{2.1}$$

The main use of the decoupling is in the following combinatorial lemma.

Lemma 2.3. *Suppose that for all $\alpha \subset [k]$ we have $|S_n(z; Y_1^\alpha, \ldots, Y_k^\alpha)| \leqslant 1$. Then*

$$\prod_{i\in[k]} \left| \sum_{j\in R_i} \left(\frac{1}{z - Z_j} - \frac{1}{z - Z_j'} \right) \right| \leqslant 2^k\,.$$

Proof. Define

$$h(Y_1, \ldots, Y_k, Y_1', \ldots, Y_k') := \sum_{\alpha\subset[k]} (-1)^{|\alpha|} S_n(z; Y_1^\alpha, \ldots, Y_k^\alpha)\,.$$

The triangle inequality shows that $|h(Y_1, \ldots, Y_k, Y_1', \ldots, Y_k')| \leqslant 2^k$. Note that

$$h(Y_1, \ldots, Y_k, Y_1', \ldots, Y_k') = \sum_{\alpha\subset[k]} (-1)^{|\alpha|} \sum_{1\leqslant i_1 < i_2 < \ldots < i_k \leqslant n} \frac{1}{z - Z_{i_1}^\alpha} \cdots \frac{1}{z - Z_{i_k}^\alpha}\,.$$

Swapping the sums, we claim that

$$\sum_{\alpha\subset[k]} (-1)^{|\alpha|} \prod_{r=1}^k \frac{1}{z - Z_{i_r}^\alpha} = \begin{cases} 0 & \text{if } \{i_1, \ldots, i_k\} \cap R_j = \varnothing \text{ for some } j; \\ \prod_{\ell=1}^k \left(\frac{1}{z - Z_{i_\ell}} - \frac{1}{z - Z_{i_\ell}'} \right) & \text{otherwise}\,. \end{cases}$$

To see this, assume first that $\{i_1, \ldots, i_k\} \cap R_j = \varnothing$ for some R_j; then when we sum over α_j, the sign changes but the quantity $\frac{1}{z - Z_{i_1}^\alpha} \cdots \frac{1}{z - Z_{i_k}^\alpha}$ does not, thus giving 0. Otherwise, we must have that each R_j contains exactly one value from $\{i_1, \ldots, i_k\}$, and thus the sum factors as stated. This shows

$$h(Y_1, \ldots, Y_k, Y_1', \ldots, Y_k') = \prod_{i\in[k]} \sum_{j\in R_i} \left(\frac{1}{z - Z_j} - \frac{1}{z - Z_j'} \right)$$

as desired. □

We now use Lemma 2.3 to identify a high-dimensional but *linear* anti-concentration event lurking in the right-hand side of (2.1).

Corollary 2.4. *Suppose $z_1, \ldots, z_L$ satisfy $|S_n(z_i, Y_1^\alpha, \ldots, Y_k^\alpha)| \leqslant 1$ for all $i \in [L]$ and $\alpha \subset [k]$. Then there is some $\ell \in [k]$ and a set I with $|I| \geqslant L/k$ so that for all $i \in I$ we have*

$$\left| \sum_{j\in R_\ell} \left(\frac{1}{z_i - Z_j} - \frac{1}{z_i - Z_j'} \right) \right| \leqslant 2\,.$$

Proof. Apply Lemma 2.3 to note that for each $i \in [L]$ we can associate some $\ell_i \in [k]$ (implicitly depend on Z_j and Z_j') for which we have

$$\left|\sum_{j \in R_{\ell_i}} \left(\frac{1}{z_i - Z_j} - \frac{1}{z_i - Z_j'}\right)\right| \leqslant 2\,.$$

Since there are only k choices for each ℓ_i, the pigeonhole principle shows that at least L/k values of i must have the same value of ℓ_i. □

With an eye towards applying Proposition 2.2, we confirm non-degeneracy of the summands appearing in Corollary 2.4:

Fact 2.5. *Suppose μ does not have finite support. Let $L \in \mathbb{N}$ and set $z_1, z_2, \ldots, z_L$ to be pairwise distinct complex numbers. Let Z and Z' be independent samples from μ. Then the vector*

$$\left(\frac{1}{z_1 - Z} - \frac{1}{z_1 - Z'}, \ldots, \frac{1}{z_L - Z} - \frac{1}{z_L - Z'}\right)$$

is non-degenerate.

Proof. The proof is similar to the case appearing in [2]. Seeking a contradiction, suppose that this random vector is degenerate, and so suppose there are (possibly complex) numbers α_j and β so that

$$\sum_{j=1}^{L} \alpha_j \left(\frac{1}{z_j - Z} - \frac{1}{z_j - Z'}\right) = \beta\,.$$

Reveal Z', and set

$$\beta' = \beta - \sum_{j=1}^{L} \frac{\alpha_j}{z_j - Z'}\,,$$

which implies that almost-surely in Z we have

$$\sum_{j=1}^{L} \frac{\alpha_j}{z_j - Z} = \beta'\,.$$

Clearing denominators, this implies that Z is the zero of a polynomial of degree at most L, which contradicts our assumption that μ does not have finite support. □

We are now ready to prove Lemma 1.3.

Proof of Lemma 1.3. By (2.1) and Corollary 2.4 there is a set R_ℓ with $|R_\ell| \geqslant \lfloor n/k \rfloor$ and a set I with $|I| \geqslant L/k$ so that

$$\begin{aligned} &\mathbb{P}\left(|S_n(z_j)| \leqslant 1 \text{ for all } j \in [L]\right) \\ &\leqslant \mathbb{P}\left(\exists\, \ell \in [k], |I| \geqslant L/k : \left|\sum_{j \in R_\ell} \left(\frac{1}{z_i - Z_j} - \frac{1}{z_i - Z_j'}\right)\right| \leqslant 2 \text{ for all } i \in I\right)^{1/2^k}. \end{aligned} \tag{2.2}$$

We now apply Proposition 2.2—noting that the non-degeneracy condition is guaranteed by Fact 2.5—to note that for each possible I and ℓ we have.

$$\mathbb{P}\left(\left|\sum_{j \in R_\ell} \left(\frac{1}{z_i - Z_j} - \frac{1}{z_i - Z_j'}\right)\right| \leqslant 2 \text{ for all } i \in I\right) \leqslant \frac{C}{(n/k)^{L/(2k)}} \tag{2.3}$$

where C depends on L and μ.

Combining (2.2) and (2.3) along with a union bound over all choices of ℓ and I shows

$$\mathbb{P}\left(|S_n(z_j)| \leqslant 1 \text{ for all } j \in [L]\right) \leqslant \left(k 2^L \frac{C}{(n/k)^{L/(2k)}}\right)^{1/2^k} .$$

Recalling $L = 2^{k+2}k$ completes the proof. □

3 Main lemmas and Proof of Theorem 1.1

Following the Angst-Malicet-Poly strategy outlined in Section 1.1, recall that we have set $P_n(z) = (z - Z_1)\cdots(z - Z_n)$ to be our random polynomial and

$$S_n(z) := \frac{P_n^{(k)}(z)}{k! P_n(z)} = \sum_{1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n} \frac{1}{z - Z_{i_1}} \cdots \frac{1}{z - Z_{i_k}} .$$

We begin with the following basic consequence of Jensen's formula, proven in [2, Prop. 2.2] .

Fact 3.1. *Let $S = Q/P$ where P and Q are two polynomials neither of which has 0 as a root. Then for each Möbius transformation $\psi \in \mathcal{M}$ we have*

$$\sum_\rho \log_- |\psi(\rho)| - \sum_\zeta \log_- |\psi(\zeta)| \leqslant \max_{z \in \psi^{-1}(C(0,1))} \log|S(z)| - \log|S(\psi^{-1}(0))|$$

where $\{\rho\}$ enumerates the roots of Q up to multiplicity and $\{\zeta\}$ enumerates the roots of P.

Recalling that Möbius transformations map circles to circles, we aim to handle the max term in Fact 3.1 first. Set $C(a, r)$ to be the circle of radius r centered at a.

Lemma 3.2. *There is a set $\mathcal{E} \subset \Omega \times \mathbb{C} \times \mathbb{R}$ with $\mathbb{P} \otimes \lambda_{\mathbb{C}} \otimes \lambda_{\mathbb{R}}(\mathcal{E}^c) = 0$ so that for all $(\omega, a, r) \in \mathcal{E}$ we have*

$$\max_{z \in C(a,r)} \log_+ |S_n(z)| = O(\log n) .$$

Proof. First consider the case $a = 0$. Note that

$$\int_{\mathbb{C}} \int_{\mathbb{R}} \frac{1}{|r - |z||^{1/2}} \mathbf{1}_{|r-|z|| \leqslant 1} \, d\lambda_{\mathbb{R}}(r) \, d\mu(z) = \int_{\mathbb{C}} 4 \, d\mu(z) = 4 < \infty$$

implying that for $\lambda_{\mathbb{R}}$-almost-every $r > 0$ we have

$$\mathbb{E}\left[\frac{1}{|r - |Z_1||^{1/2}}\right] \leqslant 1 + \mathbb{E}\left[\frac{1}{|r - |Z_1||^{1/2}} \mathbf{1}_{|r - Z| \leqslant 1}\right] < \infty . \tag{3.1}$$

Bound

$$\max_{z \in C(0,r)} |S_n(z)| \leqslant \left(\sum_{j=1}^n \frac{1}{|r - |Z_1||}\right)^k \leqslant n^k \left(\sum_{j=1}^n \frac{1}{|r - |Z_j||^{1/2}}\right)^{2k} \tag{3.2}$$

and note that for r satisfying (3.1) the strong law of large numbers implies that $\mathbb{P}$-almost-surely we have

$$\sum_{j=1}^n \frac{1}{|r - |Z_j||^{1/2}} = O(n) \tag{3.3}$$

where the implicit constant depends on the instance $\omega \in \Omega$ as well as $r \in \mathbb{R}$. Combining (3.2) with (3.3) shows that for $\lambda_{\mathbb{R}}$-almost-every $r > 0$ and $\mathbb{P}$-almost-surely we have

$$\max_{z \in C(0,r)} \log_+ |S_n(z)| = O(\log n) . \tag{3.4}$$

To show the same for arbitrary $a \in \mathbb{C}$, simply replace Z with the random variable $Z - a$ to reduce to the case of $a = 0$. This shows that for every $a \in \mathbb{C}$ there are sets Ω_a, U_a with $\mathbb{P}(\Omega_a^c) = \lambda_{\mathbb{R}}(U_a^c) = 0$ so that for all $\omega \in \Omega_a, r \in U_a$ we have that the triple (a, ω, r) satisfies the hypotheses of the Lemma. An application of the Fubini-Tonelli theorem completes the proof. □

An application of Lemma 1.3 will allow us to handle the remaining term on the right-hand side of Fact 3.1.

Lemma 3.3. *For each $k \in \mathbb{N}$, set $L = 2^{k+2}k$. Then there is a set $\mathcal{E} \subset \Omega \times \mathbb{C}^L$ with $\mathbb{P} \otimes \lambda_{\mathbb{C}}^{\otimes L}(\mathcal{E}^c) = 0$ so that for all $(\omega, z_1, \ldots, z_L) \in \mathcal{E}$ and all n large enough there is at least one $j \in [L]$ so that $|S_n(z_j)| \geqslant 1$ (with the convention $|S_n(z)| = +\infty$ if z is a pole of S_n).*

Proof. For each L-tuple $(z_1, \ldots, z_L)$ of pairwise distinct points, Lemma 1.3 and the Borel-Cantelli lemma show that almost surely for sufficiently large n we have $|S_n(z_j)| \geqslant 1$ for at least one $j \in [L]$. Since the set of pairwise distinct L-tuples has complement of measure 0 under the measure $\lambda_{\mathbb{C}}^{\otimes L}$, an application of the Fubini-Tonelli theorem completes the proof. □

Finally, the strong law of large numbers will control the sum over $\{\zeta\}$:

Fact 3.4. *There is a set $\mathcal{E} \subset \Omega \times \mathcal{M}$ with $\mathbb{P} \otimes \lambda_{\mathcal{M}}(\mathcal{E}^c) = 0$ so that for all $(\omega, \psi) \in \mathcal{E}$ we have*

$$\lim_{n\to\infty} \int_{\mathbb{C}} \log_- |\psi(z)| \, d\mu_n(z) = \int_{\mathbb{C}} \log_- |\psi(z)| \, d\mu(z) \, .$$

Proof. By definition, we have $\int_{\mathbb{C}} \log_- |\psi(z)| d\mu_n(z) = \frac{1}{n} \sum_{k=1}^{n} \log_- |\psi(Z_k(z))| \, d\mu(z)$. By the strong law of large numbers, we have that almost surely

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log_- |\psi(Z_k(z))| = \int_{\mathbb{C}} \log_- |\psi(z)| \, d\mu(z) \, . \qquad \square$$

We are now ready to verify the assumptions of Lemma 1.2 for cluster points of our sequences of random measures.

Lemma 3.5. *Suppose μ does not have finite support. There is a set $\mathcal{E} \subset \Omega \times \mathcal{M}$ with $\mathbb{P} \otimes \lambda_{\mathcal{M}}(\mathcal{E}^c) = 0$ so that the following holds. For every cluster point $\widehat{\nu}_\infty$ of the sequence of empirical probability measures $(\nu_n^{(k)})_{n\geqslant 1}$ we have*

$$\int_{\widehat{\mathbb{C}}} \log_- |\psi(z)| \, d\widehat{\nu}_\infty(z) \leqslant \int_{\mathbb{C}} \log_- |\psi(z)| \, d\mu(z)$$

for all $(\omega, \psi) \in \mathcal{E}$.

Proof. By combining Lemma 3.2, Lemma 3.3, and Fact 3.4 we see that there is a set Ω_1 with $\mathbb{P}(\Omega_1) = 1$ for which the following holds: for $\lambda_{\mathcal{M}}^{\otimes L}$-almost-every tuple of Möbius transformations $(\psi_1, \ldots, \psi_L)$ we have

1. $\max_{z \in \psi_j^{-1}(C(0,1))} \log |S_n(z)| = O(\log n)$ for all $j \in [L]$.
2. For each n large enough, there is some $j \in [L]$ so that $\log |S_n(\psi_j^{-1}(0)| \geqslant 0$.
3. $\lim_{n\to\infty} \int_{\mathbb{C}} \log_- |\psi_j(z)| \, d\mu_n(z) = \int_{\mathbb{C}} \log_- |\psi_j(z)| \, d\mu(z)$ for all $j \in [L]$.

We note that for Item 2 we use the fact that for $\lambda_{\mathcal{M}}^{\otimes L}$-almost-every tuple $(\psi_j)_{j=1}^L$, the values $(\psi_j^{-1}(0))_{j=1}^L$ are pairwise distinct. Fix an instance $\omega \in \Omega_1$ and a tuple $(\psi_1, \ldots, \psi_L)$ for which the above three items hold. Combining Jensen's formula Fact 3.1 along with

Item 1 and Item 2, we see that for each n sufficiently large there is some $j \in L$ for which we have

$$\left(1-\frac{k}{n}\right)\int_{\mathbb{C}} \log_-|\psi_j(z)| d\nu_n^{(k)}(z) \leqslant \int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\mu_n(z) + O\left(\frac{\log n}{n}\right). \tag{3.5}$$

By Item 3 we have that

$$\lim_{n\to\infty}\int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\mu_n(z) = \int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\mu(z). \tag{3.6}$$

For a cluster point $\widehat{\nu}_\infty$, there exists a subsequence $(\nu_{n_i})_{i\geqslant 1}$ converging to $\widehat{\nu}_\infty$. By truncating the $\log_-$ and applying the monotone convergence theorem we see that for each $j \in [L]$ we have

$$\int_{\widehat{\mathbb{C}}} \log_-|\psi_j(z)|\, d\widehat{\nu}_\infty(z) \leqslant \limsup_{i\to\infty}\int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\nu_{n_i}^{(k)}(z). \tag{3.7}$$

Thinning our subsequence further so that Item 2 holds for a single $j \in [L]$ for all $(n_i)_{i\geqslant 1}$, we combine (3.5), (3.6) and (3.7) to obtain

$$\int_{\widehat{\mathbb{C}}} \log_-|\psi_j(z)|\, d\widehat{\nu}_\infty(z) \leqslant \int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\mu(z). \tag{3.8}$$

Defining the set $E \subset \mathcal{M}$ (depending on the instance $\omega \in \Omega_1$) via

$$E := \left\{\psi \in \mathcal{M} : \exists \text{ cluster point } \widehat{\nu}_\infty \text{ with } \int_{\widehat{\mathbb{C}}} \log_-|\psi_j(z)|\, d\widehat{\nu}_\infty(z) > \int_{\mathbb{C}} \log_-|\psi_j(z)|\, d\mu(z)\right\}$$

we see that (3.8) shows that $\lambda_{\mathcal{M}}^{\otimes L}(E^L) = 0$ where E^L is the L-fold cartesian product of E with itself; this implies that $\lambda_{\mathcal{M}}(E) = 0$ as desired. □

Proof of Theorem 1.1. First note that if μ has finite support, then the theorem follows immediately by the strong law of large numbers. As such, we will assume throughout that μ does not have finite support. By Lemma 3.5, there is a set Ω_1 with $\mathbb{P}(\Omega_1) = 1$ so that for each $\omega \in \Omega_1$, for $\lambda_{\mathcal{M}}$-almost-all $\psi \in \mathcal{M}$ we have

$$\int_{\widehat{\mathbb{C}}} \log_-|\psi(z)|\, d\widehat{\nu}_\infty(z) \leqslant \int_{\mathbb{C}} \log_-|\psi(z)|\, d\mu(z).$$

By Lemma 1.2, this implies that each such cluster point satisfies $\widehat{\nu}_\infty = \mu$. This shows that each subsequence of $\{\nu_n^{(k)}\}_{n\geqslant 1}$ contains a further subsequence that converges to μ, thus showing that for each $\omega \in \Omega_1$ we have ν_n converges to μ. □

References

[1] T. Alazard, O. Lazar, and Q. H. Nguyen. On the dynamics of the roots of polynomials under differentiation. *Journal de Mathématiques Pures et Appliquées*, 162:1–22, 2022. MR4417282

[2] J. Angst, D. Malicet, and G. Poly. Almost sure behavior of the critical points of random polynomials. *Bulletin of the London Mathematical Society*, 56(2):767–782, 2024. MR4711583

[3] S.-S. Byun, J. Lee, and T. R. Reddy. Zeros of random polynomials and their higher derivatives. *Trans. Amer. Math. Soc.*, 375(9):6311–6335, 2022. MR4474893

[4] P. Cheung, T. Ng, and S. Yam. Critical points of random finite blaschke products with independent and identically distributed zeros. *Complex Analysis and Potential Theory with Applications*, 2014.

[5] P.-L. Cheung, T. W. Ng, J. Tsai, and S. C. P. Yam. Higher-order, polar and Sz.-Nagy's generalized derivatives of random polynomials with independent and identically distributed zeros on the unit circle. *Comput. Methods Funct. Theory*, 15(1):159–186, 2015. MR3318313

[6] K. P. Costello. Bilinear and quadratic variants on the Littlewood-Offord problem. *Israel J. Math.*, 194(1):359–394, 2013. MR3047075

[7] K. P. Costello, T. Tao, and V. Vu. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006. MR2267289

[8] J. Hoskins and Z. Kabluchko. Dynamics of zeroes under repeated differentiation. *arXiv preprint arXiv:2010.14320*, 2020. MR4669281

[9] J. G. Hoskins and S. Steinerberger. A semicircle law for derivatives of random polynomials. *International Mathematics Research Notices*, 2022(13):9784–9809, 2022. MR4447137

[10] Z. Kabluchko. Critical points of random polynomials with independent identically distributed roots. *Proceedings of the American Mathematical Society*, 143(2):695–702, 2015. MR3283656

[11] Z. Kabluchko. Repeated differentiation and free unitary Poisson process. *arXiv preprint arXiv:2112.14729*, 2021.

[12] A. Kiselev and C. Tan. The flow of polynomial roots under differentiation. *Annals of PDE*, 8(2):1–69, 2022. MR4458083

[13] M. Kwan and L. Sauermann. An algebraic inverse theorem for the quadratic Littlewood-Offord problem, and an application to Ramsey graphs. *Discrete Anal.*, pages Paper No. 12, 34, 2020. MR4143152

[14] M. Kwan and L. Sauermann. Resolution of the quadratic Littlewood–Offord problem. *arXiv preprint arXiv:2312.13826*, 2023.

[15] R. Meka, O. Nguyen, and V. Vu. Anti-concentration for polynomials of independent random variables. *Theory of Computing*, 12(1):1–17, 2016. MR3542863

[16] M. Michelen and X.-T. Vu. Zeros of a growing number of derivatives of random polynomials with independent roots. *arXiv preprint arXiv:2212.11867*, 2022. MR4741259

[17] H. H. Nguyen and V. H. Vu. Small ball probability, inverse theorems, and applications. In *Erdös centennial*, volume 25 of *Bolyai Soc. Math. Stud.*, pages 409–463. János Bolyai Math. Soc., Budapest, 2013. MR3203607

[18] S. O'Rourke and S. Steinerberger. A nonlocal transport equation modeling complex roots of polynomials under differentiation. *Proceedings of the American Mathematical Society*, 149(4):1581–1592, 2021. MR4242313

[19] R. Pemantle and I. Rivin. The distribution of zeros of the derivative of a random polynomial. In *Advances in Combinatorics: Waterloo Workshop in Computer Algebra, W80, May 26-29, 2011*, pages 259–273. Springer, 2013. MR3363974

[20] S. Steinerberger. Free convolution of measures via roots of polynomials. *arXiv preprint arXiv:2009.03869*, 2020. MR4669280

[21] S. D. Subramanian. On the distribution of critical points of a polynomial. *Electron. Commun. Probab.*, 17:no. 37, 9, 2012. MR2970701

Acknowledgments. Both authors are supported in part by NSF grants DMS-2137623 and DMS-2246624.