

ON THE DISTRIBUTION OF CRITICAL POINTS OF A POLYNOMIAL

SNEHA DEY SUBRAMANIAN

ABSTRACT. This paper proves that if points Z_1, Z_2, \dots are chosen independently and identically using some measure μ from the unit circle in the complex plane, with $p_n(z) = (z - Z_1)(z - Z_2)\dots(z - Z_n)$, then the empirical distribution of the critical points of p_n converges weakly to μ .

1. INTRODUCTION

Across many fields of mathematics, one of the fundamental questions about a function is the location of its zeros. Perhaps the most famous example is the Riemann Hypothesis. Beyond that, entire fields such as algebraic geometry and the emergent study of stable functions have locations of zeros as their focus.

The relation between the zeros of a function and the zeros of its derivative (the critical points) is interesting and not always obvious. In the case where all zeros are real, Rolle's theorem tells us that the zeros of the derivative interlace the zeros of the function itself. In the case of complex polynomials the analogous result is the Gauss-Lucas theorem which states that the zeros of the derivative of f must lie in the convex hull of the zeros of f and gives a representation of the zeros of f' as convex combinations of the zeros of f . A corollary of this is that differentiating preserves stability. Differentiation is also known never to increase the number of non-real zeros of a polynomial.

Two famous conjectures in this area are the conjectures of Sendov and Smale. The former, made by Blagovest Sendov during the 1950's, states that if the roots z_1, z_2, \dots, z_n of a polynomial all lie inside the closed unit disc, then for each root of the polynomial, the closed unit disc centered at the root must contain at least one critical point. The latter, made by Steve Smale, states that if f is a polynomial of

Ph.D. candidate, Department of Mathematics, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA-19104, U.S.A. Email: ssneha@math.upenn.edu.

degree n with at least one root 0 and $f'(0) \neq 0$, then,

$$\min \left\{ \frac{|f(\xi)|}{|\xi||f'(0)|} : f'(\xi) = 0 \right\} \leq K,$$

where $K = 1$ or $\frac{n-1}{n}$. Sendov's conjecture has been proven for the case when z_1, z_2, \dots, z_n all lie on the unit circle, whereas Smale's conjecture has been proven for when f has all its roots, save 0, on the unit circle. The most general forms of these conjectures are still unsolved. More information on these conjectures and proofs of some of the special cases can be found in [RS02].

A probabilistic study on the roots of derivatives of polynomials was done by Pemantle and Rivin in [PR12]. Let f be a polynomial with n roots that are chosen independently and uniformly from a measure μ on the complex plane. They conjectured that the empirical distribution of the roots of f' converges weakly to μ as $n \rightarrow \infty$. They prove this in the special case when μ has finite 1-energy, namely when μ satisfies

$$\int \int \frac{1}{|z-w|} d\mu(z) d\mu(w) < \infty.$$

This condition cannot hold, however, when μ is supported on any set of dimension 1 or less. The aim of the present paper is to extend their result to the case of any measure supported on the unit circle.

2. NOTATIONS AND BACKGROUND

Say, Z_1, Z_2, \dots is a sequence of points chosen i.i.d. with respect to some distribution μ on the unit circle. Write, $Z_k = \exp(2\pi i\theta_k)$, so that $\{\theta_k\}$ is a collection of IID random variables whose common law is supported on $[0, 1]$, which we denote by ν .

Let

$$p_n(z) = (z - Z_1)(z - Z_2)\dots(z - Z_n),$$

and $y_1^{(n)}, y_2^{(n)}, \dots, y_{n-1}^{(n)}$ be the roots of $p_n'(z)$.

For $k \geq 1$, let $c_k = \mathbf{E}(Z^k)$, where $Z \sim \mu$. We know, by the Strong Law of Large Numbers,

$$\frac{Z_1^k + Z_2^k + \dots + Z_n^k}{n} \xrightarrow{a.s.} c_k.$$

Denote by $\mathcal{Z}(f)$ the empirical distribution of the roots of a random polynomial f . That is, if f has roots X_1, X_2, \dots, X_m , then $\mathcal{Z}(f) = \frac{1}{m} \sum_{j=1}^m \delta_{X_j}$.

We shall write \mathbb{D} for the open unit disc, and \mathcal{C} for the unit circle.

In their paper, [PR12], the authors conjectured that, for any distribution μ on the closed unit disc, $\mathcal{Z}(p'_n)$ converges weakly to μ . That paper also proves the following proposition.

Proposition 2.1. *Let μ be the uniform measure on \mathcal{C} . Then $\mathcal{Z}(p'_n)$ converges to \mathcal{C} in probability, that is, $P(\mathcal{Z}(S) \geq \epsilon) \rightarrow 0$ for any $\epsilon > 0$ and any closed set $S \subset \mathbb{D}$, disjoint from \mathcal{C} . \square*

In this note, we shall generalize this to prove that

Lemma 2.2. *For any distribution μ on \mathcal{C} , $\mathcal{Z}(p'_n)$ converges to \mathcal{C} in probability. In fact, if μ is not uniform on \mathcal{C} , the convergence is almost everywhere.*

The above leads us to prove our main result, which is a special case of the aforementioned conjecture in [PR12]:

Theorem 2.3. *For any distribution μ on \mathcal{C} , $\mathcal{Z}(p'_n)$ converges weakly to μ on \mathcal{C} .*

The proof, as shall be seen in forthcoming sections, can be divided in to two parts, the latter following a pattern similar to the proof of Weyl's equidistribution criterion (see, for example [Ch68]). The former requires the following theorem proved in [CN06] regarding a companion matrix of the critical points.

Proposition. *If $z_1, z_2, \dots, z_n \in \mathbb{C}$, and y_1, y_2, \dots, y_{n-1} are the critical points of the polynomial $p_n(z) = (z - z_1)(z - z_2) \dots (z - z_n)$, then, the matrix*

$$(1) \quad D \left(I - \frac{J}{n} \right) + \frac{z_n}{n} J$$

has y_1, y_2, \dots, y_{n-1} as its eigenvalues, where $D = \text{diag}(z_1, z_2, \dots, z_{n-1})$, I is the identity matrix of order $n - 1$ and J is the $(n - 1) \times (n - 1)$ matrix of all entries 1. \square

3. PROOFS OF LEMMA 2.2 AND THEOREM 2.3

We first begin by proving a small lemma.

Lemma 3.1. *Let μ be a distribution on the unit circle \mathcal{C} with $c_k = \mathbf{E}(Z^k)$, where $Z \sim \mu$. Then $c_k = 0$ for all $k \geq 1$ if and only if μ is uniform on \mathcal{C} .*

Proof. Clearly if μ is uniform on \mathcal{C} then $c_k = 0$ for all $k \geq 1$. Now say μ is not uniform on the circle but we still have $c_k = 0$ for all $k \geq 1$. Then the law ν is not uniform on $[0, 1]$. Now, if Z_1, Z_2, \dots are points on \mathcal{C} , chosen i.i.d. using μ , and if we write $Z_j = \exp(2\pi i\theta_j)$, $j = 1, 2, \dots$, then $\theta_1, \theta_2, \dots$ are points in $[0, 1]$ that are i.i.d. ν .

As mentioned earlier, by the Strong Law of Large Numbers, for all $k \geq 1$,

$$\frac{Z_1^k + Z_2^k + \dots + Z_n^k}{n} \xrightarrow{a.s.} 0,$$

and so by Weyl's criterion, for any $0 \leq a < b \leq 1$,

$$\frac{\sum_{j=1}^n 1_{\{\theta_j \in [a, b]\}}}{n} \xrightarrow{a.s.} b - a.$$

But $1_{\{\theta_j \in [a, b]\}}$, $j = 1, 2, \dots$ are i.i.d. random variables taking values 0 or 1 with expectation $\nu([a, b])$. So by Strong Law of Large Numbers,

$$\frac{\sum_{j=1}^n 1_{\{\theta_j \in [a, b]\}}}{n} \xrightarrow{a.s.} \nu([a, b]).$$

Since, ν is not uniform on $[0, 1]$ we have arrived at a contradiction. Therefore, there must exist at least one non-zero c_k . \square

We proceed to use this fact for the proof of Lemma 2.2.

Proof of Lemma 2.2. Assume μ is not the uniform distribution on the circle (as the uniform case has been taken care of in [PR12]). Then, as mentioned above, there is at least one non-zero c_k . Thus the power series function $f(z) = \sum_{k=0}^{\infty} \bar{c}_{k+1} z^k$ exists at every point $z \in \mathbb{D}$, is analytic there (since $|c_k| < 1, \forall k$), and so has only finitely many zeros inside any r -ball, where $r < 1$.

Define

$$V_n(z) = \frac{p'_n(z)}{np_n(z)} = \frac{1}{n} \sum_{j=1}^n \frac{1}{z - Z_j}.$$

V_n has $n - 1$ zeros, which are exactly the zeros of $p'_n(z)$, and n poles, which are exactly the zeros of $p_n(z)$. Thus $V_n(z)$ is analytic inside \mathbb{D} . We shall show that as $n \rightarrow \infty$, V_n converges inside the disc to $-f$, uniformly over compact sets. To see

this, note that for $z \in \mathbb{D}$,

$$\begin{aligned} V_n(z) &= \frac{1}{n} \sum_{j=1}^n \frac{-1/Z_j}{1 - z/Z_j} = -\frac{1}{n} \sum_{j=1}^n \frac{1}{Z_j} \left(1 + \frac{z}{Z_j} + \frac{z^2}{Z_j^2} + \dots \right) \\ &= -\frac{1}{n} \sum_{j=1}^n \sum_{k=0}^{\infty} \bar{Z}_j^{k+1} z^k = -\sum_{k=0}^{\infty} \bar{a}_n^{k+1} z^k, \end{aligned}$$

where, we write \bar{a}_n^{k+1} for the k th power sum average $\frac{Z_1^k + Z_2^k + \dots + Z_n^k}{n}$. As stated earlier, $\bar{a}_n^k \xrightarrow{a.s.} c_k$ for all $k \geq 1$.

Let $0 < r < 1$. Given any $\delta > 0$, $\exists K \geq 1$ such that

$$\sum_{k=K}^{\infty} r^k = \frac{r^K}{1-r} < \frac{\delta}{4}.$$

Corresponding to the chosen K , there exists an $N \geq 1$ such that,

$$|a_n^k - c_k| < \frac{\delta(1-r)}{2},$$

$\forall n \geq N$ and $\forall k = 1, 2, \dots, K-1$. Therefore, $\forall n \geq N$ and all $z \in B_r(0)$,

$$\begin{aligned} |V_n(z) + f(z)| &\leq \sum_{k=0}^{K-1} |a_n^k - c_k| r^k + \sum_{k=K+1}^{\infty} |a_n^k - c_k| r^k \\ &\leq \frac{\delta(1-r)}{2} \cdot (1 + r + r^2 + \dots + r^{K-1}) + 2 \cdot \frac{\delta}{4} < \delta, \end{aligned}$$

which proves uniform convergence of V_n to $-f$ over compact sets.

Hurwitz's theorem (see, for example, [Co78]) says that, if g_n is a sequence of analytic functions that converges uniformly on compact subsets of an open set G in the complex plane to an analytic function g , then, for any closed disc D contained in G , if g has no zeros on the boundary of D , there shall exist a natural number N for which g_n and g have the same number of zeros in the interior of D , for all $n \geq N$.

Using Hurwitz's theorem, given any $0 < r < 1$, there exists an $M \geq 1$ for which V_n and f have the same number of zeros inside $B_r(0)$ for all $n \geq M$. That is, p'_n and f shall have the same number of zeros inside $B_r(0)$ for all $n \geq M$. But, as discussed above, f has only finitely many zeros inside $B_r(0)$. Thus $\mathcal{Z}(p'_n)$ converges to the unit circle almost surely. \square

Our main result, Theorem 2.3, will be a consequence of the following proposition together with Lemma 3.3 below.

Proposition 3.2. *Given any sequence of points z_1, z_2, \dots with $|z_n| \leq M$ for all n , and $\frac{z_1^k + z_2^k + \dots + z_n^k}{n} \rightarrow c_k$ as $n \rightarrow \infty$, $\forall k \geq 1$, the critical points $y_1^{(n)}, y_2^{(n)}, \dots, y_{n-1}^{(n)}$ of $p_n(z) = (z - z_1)(z - z_2)\dots(z - z_n)$ also satisfy*

$$\frac{(y_1^{(n)})^k + (y_2^{(n)})^k + \dots + (y_{n-1}^{(n)})^k}{n-1} \rightarrow c_k \text{ as } n \rightarrow \infty,$$

$\forall k \geq 1$.

Proof. Note that, it is easy to see that this theorem holds true for $k = 1$, because the average of the critical points is exactly equal to the average of the roots (by comparing the coefficients of z^{n-1} in $p_n(z)$ with z^{n-2} of $p'_n(z)$). To prove the result for general k , we use a result of [CN06], mentioned as a proposition in Section 2, to see that for $k \geq 2$, $(y_1^{(n)})^k, (y_2^{(n)})^k, \dots, (y_{n-1}^{(n)})^k$ are the eigenvalues of $[D(I - \frac{1}{n}J) + \frac{z_n}{n}J]^k$, and so,

$$(y_1^{(n)})^k + (y_2^{(n)})^k + \dots + (y_{n-1}^{(n)})^k = \text{Tr} \left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^k.$$

Note that the expansion of $[D(I - \frac{1}{n}J) + \frac{z_n}{n}J]^k$ is the sum of all terms such as

$$(2) \quad D^{l_1} \left(-\frac{DJ}{n} \right)^{l_2} \left(\frac{z_n J}{n} \right)^{l_3} D^{l_4} \left(-\frac{DJ}{n} \right)^{l_5} \left(\frac{z_n J}{n} \right)^{l_6} \dots D^{l_{3k-2}} \left(-\frac{DJ}{n} \right)^{l_{3k-1}} \left(\frac{z_n J}{n} \right)^{l_{3k}}$$

where the exponents l_1, l_2, \dots, l_{3k} are non-zero integers, with $l_{3j-2} + l_{3j-1} + l_{3j} = 1$ for all $j = 1, 2, \dots, k$. Clearly the number of such terms is 3^k , which does not depend on n , and so, if we find that the trace of the matrix in the expression (2) converges as $n \rightarrow \infty$ to $a_{l_1, l_2, \dots, l_{3k}}$, then the trace of $[D(I - \frac{1}{n}J) + \frac{z_n}{n}J]^k$ converges to $\sum a_{l_1, l_2, \dots, l_{3k}}$.

Henceforth, we fix l_1, l_2, \dots, l_{3k} . Now, note that $J^m = (n-1)^{m-1}J^{m-1}$ for any $m \geq 1$, and

$$\begin{aligned} (D^p J)(D^q J) &= \begin{pmatrix} z_1^p & z_1^p & \dots & z_1^p \\ z_2^p & z_2^p & \dots & z_2^p \\ \vdots & \vdots & \ddots & \vdots \\ z_{n-1}^p & z_{n-1}^p & \dots & z_{n-1}^p \end{pmatrix} \cdot \begin{pmatrix} z_1^q & z_1^q & \dots & z_1^q \\ z_2^q & z_2^q & \dots & z_2^q \\ \vdots & \vdots & \ddots & \vdots \\ z_{n-1}^q & z_{n-1}^q & \dots & z_{n-1}^q \end{pmatrix} \\ &= \left(\sum_{i=1}^{n-1} z_i^q \right) (D^p J), \end{aligned}$$

for any $p, q \geq 1$.

The above tells us that there exists $p, q, s_0, s_1, s_2, \dots, s_{k-1} \geq 0$ such that, term (2) is of the form

$$(3) \quad (-1)^p \cdot z_n^q \cdot \left(\frac{n-1}{n}\right)^{s_0} \cdot \left(\frac{\sum_{i=1}^{n-1} z_i}{n}\right)^{s_1} \cdot \left(\frac{\sum_{i=1}^{n-1} z_i^2}{n}\right)^{s_2} \cdot \dots \cdot \left(\frac{\sum_{i=1}^{n-1} z_i^{k-1}}{n}\right)^{s_{k-1}} \cdot M,$$

where the numbers $p, q, s_0, s_1, \dots, s_{k-1}$ are determined solely by the l_i 's (and so, are independent of n). For instance,

$$p = l_2 + l_5 + \dots + l_{3k-1}, \text{ and} \\ q = l_3 + l_6 + \dots + l_{3k}.$$

Also, M can only be one of the following terms: D^k or $\frac{D^m J}{n}$ or $\frac{D^{m_1} J}{n} D^{m_2}$ for some $m, m_1, m_2 \geq 0$, which are fixed, $\leq k$, and dependent only on the l_i 's (this fact, and also the scalar coefficient in (3), can be checked by simply using induction on k). Furthermore, the scalar coefficient in (3) is always $O(1)$.

Observe that, if $M = D^k$, then the scalar coefficient of M is equal to 1 (since there is only one D^k term in the expansion of $[D(I - \frac{1}{n}J) + \frac{z_n}{n}J]^k$) and $\frac{Tr(M)}{n} \rightarrow c_k$.

On the other hand, if $M = \frac{D^m J}{n}$, then

$$Tr(M) = \frac{(y_1^{(n)})^m + (y_2^{(n)})^m + \dots + (y_{n-1}^{(n)})^m}{n} = o(n),$$

and if $M = \frac{D^{m_1} J}{n} D^{m_2}$,

$$Tr(M) = Tr\left(D^{m_1+m_2} \frac{J}{n}\right) \\ = \frac{(y_1^{(n)})^{m_1+m_2} + (y_2^{(n)})^{m_1+m_2} + \dots + (y_{n-1}^{(n)})^{m_1+m_2}}{n} = o(n).$$

Thus,

$$\frac{Tr\left[D\left(I - \frac{1}{n}J\right) + \frac{z_n}{n}J\right]^k}{n} \rightarrow c_k \text{ as } n \rightarrow \infty.$$

□

We require now a simple lemma about the convergence of weighted averages with some suitable conditions on the weights.

Lemma 3.3. *Let $\{x_{nk}\}_{n,k=1}^\infty$ be a double sequence of real numbers or real-valued random variables, for which, $x_{nk} \in (0, 1)$ for all n and for all k . Also, let $\{b_{nk}\}$ be*

a double sequence of complex-valued random variables, with $|b_{nk}| \leq M$ for all n and for all k . Then, if

$$\frac{x_{1n} + x_{2n} + \dots + x_{nn}}{n} \rightarrow 0,$$

in probability, almost surely, or pointwise,

$$\frac{x_{1n}b_{1n} + x_{2n}b_{2n} + \dots + x_{nn}b_{nn}}{n} \rightarrow 0,$$

respectively in probability, almost surely, or pointwise.

Proof. We have,

$$\begin{aligned} \left| \frac{x_{1n}b_{1n} + x_{2n}b_{2n} + \dots + x_{nn}b_{nn}}{n} \right| &\leq \frac{x_{1n}|b_{1n}| + x_{2n}|b_{2n}| + \dots + x_{nn}|b_{nn}|}{n} \\ &\leq M \left(\frac{x_{1n} + x_{2n} + \dots + x_{nn}}{n} \right), \end{aligned}$$

which gives us the result. \square

We now have all the tools required to prove our main result, namely Theorem 2.3.

Proof of Theorem 2.3. Say we write,

$$y_j^{(n)} = r_j^{(n)} \exp(2\pi i \phi_j^{(n)}), j = 1, 2, \dots, n-1.$$

The proof will consist of three major segments. Our first task is to prove that

$$\frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \xrightarrow{P} 1.$$

In fact, unless μ is uniform on the circle, we will show that

$$\frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \xrightarrow{a.s.} 1.$$

Next, we shall use the above information and Lemma 3.3 to show that

$$\frac{\exp(2k\pi i \phi_1^{(n)}) + \exp(2k\pi i \phi_2^{(n)}) + \dots + \exp(2k\pi i \phi_{n-1}^{(n)})}{n-1} \xrightarrow{P} c_k.$$

(Again, the convergence is almost sure, unless μ is uniform on \mathcal{C} .)

Finally, using arguments analogous to those in the proof of Weyl's equidistribution criterion, we shall arrive at our final result.

Assume, initially, that μ is not the uniform law on \mathcal{C} . For the first task as noted above, observe that, by Lemma 2.2, given any $\epsilon > 0$,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} 1_{\{r_j^{(n)} \in [1-\epsilon, 1]\}} \xrightarrow{a.s.} 1.$$

Now, for any fixed positive integer k , $(1-\epsilon)^k 1_{\{r_j^{(n)} \in [1-\epsilon, 1]\}} \leq (r_j^{(n)})^k \leq 1$, and so

$$(4) \quad (1-\epsilon)^k \cdot \frac{1}{n-1} \sum_{j=1}^{n-1} 1_{\{r_j^{(n)} \in [1-\epsilon, 1]\}} \leq \frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \leq 1.$$

Clearly then, a simple squeeze theorem argument gives us

$$(5) \quad \frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \xrightarrow{a.s.} 1.$$

Now, from Proposition 3.2, for any positive integer k ,

$$\begin{aligned} & \frac{(y_1^{(n)})^k + (y_2^{(n)})^k + \dots + (y_{n-1}^{(n)})^k}{n-1} \xrightarrow{a.s.} c_k, \\ \implies & \frac{(r_1^{(n)})^k \exp(2k\pi i \phi_1^{(n)}) + (r_2^{(n)})^k \exp(2k\pi i \phi_2^{(n)}) + \dots + (r_{n-1}^{(n)})^k \exp(2k\pi i \phi_{n-1}^{(n)})}{n-1} \xrightarrow{a.s.} c_k. \end{aligned}$$

Then, applying Lemma 3.3 with $\{x_{nm} = 1 - r_n^{(m)}\}$, we get

$$(6) \quad \frac{\exp(2k\pi i \phi_1^{(n)}) + \exp(2k\pi i \phi_2^{(n)}) + \dots + \exp(2k\pi i \phi_{n-1}^{(n)})}{n-1} \xrightarrow{a.s.} c_k.$$

Now, for the final stage of our proof,

$$\begin{aligned} c_k &= \mathbf{E}(Z^k), \text{ where, } Z \sim \mu. \\ \implies c_k &= \mathbf{E}(\exp(2k\pi i \Theta)) = \mathbf{E}(\cos(2k\pi \Theta)) + i\mathbf{E}(\sin(2k\pi \Theta)), \text{ where, } \Theta \sim \nu. \end{aligned}$$

So, (6) gives,

$$\begin{aligned} & \frac{\cos(2k\pi \phi_1^{(n)}) + \cos(2k\pi \phi_2^{(n)}) + \dots + \cos(2k\pi \phi_{n-1}^{(n)})}{n} \xrightarrow{a.s.} \mathbf{E}(\cos(2k\pi \Theta)), \\ & \frac{\sin(2k\pi \phi_1^{(n)}) + \sin(2k\pi \phi_2^{(n)}) + \dots + \sin(2k\pi \phi_{n-1}^{(n)})}{n} \xrightarrow{a.s.} \mathbf{E}(\sin(2k\pi \Theta)). \end{aligned}$$

Then, for any trigonometric polynomial,

$$q(x) = a_0 + \sum_{j=1}^r (a_j \cos(2\pi j x) + b_j \sin(2\pi j x)), \quad a_j, b_j \in \mathbb{R}, \quad j = 1, 2, \dots, r,$$

we have,

$$(7) \quad \frac{\sum_{j=1}^{n-1} q(\phi_j^{(n)})}{n} \xrightarrow{a.s.} \mathbf{E}(q(\Theta)).$$

Let f be a continuous real-valued function on $[0, 1]$ and fix $\epsilon > 0$. Stone-Weierstrass theorem ([St48]) states that for any compact Hausdorff space X , and any subalgebra S of the space $C(X, \mathbb{R})$ (continuous functions from X to \mathbb{R}) that separates points and contains at least one non-zero function, S will be dense in $C(X, \mathbb{R})$ with respect to uniform norm. In particular, this means that trigonometric polynomials are dense in the space of continuous functions on $[0, 1]$ with the uniform norm. Hence, there exists a trigonometric polynomial q such that $|f - q| < \epsilon$. So

$$\begin{aligned} \left| \frac{\sum_{j=1}^{n-1} f(\phi_j^{(n)})}{n} - \mathbf{E}(f(\Theta)) \right| &\leq \left| \frac{\sum_{j=1}^{n-1} f(\phi_j^{(n)})}{n} - \frac{\sum_{j=1}^{n-1} q(\phi_j^{(n)})}{n} \right| \\ &\quad + \left| \frac{\sum_{j=1}^{n-1} q(\phi_j^{(n)})}{n} - \mathbf{E}(q(\Theta)) \right| + |\mathbf{E}(q(\Theta)) - \mathbf{E}(f(\Theta))| \\ &\leq \frac{\sum_{j=1}^{n-1} |f(\phi_j^{(n)}) - q(\phi_j^{(n)})|}{n} \\ &\quad + \left| \frac{\sum_{j=1}^{n-1} q(\phi_j^{(n)})}{n} - \mathbf{E}(q(\Theta)) \right| + \mathbf{E}|q(\Theta) - f(\Theta)|. \end{aligned}$$

The first and third terms on the right hand side are each $< \epsilon$ while the second term goes to 0 almost surely, by (7). Hence for any f continuous on $[0, 1]$,

$$(8) \quad \frac{\sum_{j=1}^{n-1} f(\phi_j^{(n)})}{n} \xrightarrow{a.s.} \mathbf{E}(f(\Theta)),$$

and this holds for complex-valued continuous functions as well (which is easily seen by comparing the real and imaginary parts). Thus, the joint empirical distribution of $\phi_j^{(n)}, j = 1, 2, \dots, n-1$, converges weakly to ν , which means that the joint empirical distribution of $\exp(2\pi i \phi_j^{(n)}), j = 1, 2, \dots, n-1$, converges weakly to μ . This, along with Lemma 2.2, gives us the desired result for μ not uniform on \mathcal{C} .

Now suppose μ is the uniform law on the unit circle. Then, by (4), for any $\delta > 0$,

$$\begin{aligned} \mathbf{P} \left(1 - \frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \geq \delta \right) &\leq \mathbf{P} \left(1 - (1-\epsilon)^k \cdot \frac{1}{n-1} \sum_{j=1}^{n-1} 1_{\{r_j^{(n)} \in [1-\epsilon, 1]\}} \geq \delta \right) \\ &\leq \mathbf{P} \left(1 - \frac{1}{n-1} \sum_{j=1}^{n-1} 1_{\{r_j^{(n)} \in [1-\epsilon, 1]\}} \geq \frac{\delta - 1 + (1-\epsilon)^k}{(1-\epsilon)^k} \right). \end{aligned}$$

We may choose ϵ in such a way that $1 - (1-\epsilon)^k < \delta/2$, and so, the right hand side of the above inequality goes to 0. Thus, for any positive integer k ,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} (r_j^{(n)})^k \xrightarrow{P} 1.$$

Note that the above is a slightly weaker version of (5), since the convergence is now in probability, and not almost sure.

For the rest of the proof, observe that we can follow the same arguments as in the non-uniform case, except that the almost sure convergence in each of the statements will be replaced by convergence in probability. Thus, we shall arrive at

$$\frac{\sum_{j=1}^{n-1} f(\phi_j^{(n)})}{n} \xrightarrow{P} \mathbf{E}(f(\Theta)),$$

for any continuous function $f : [0, 1] \rightarrow \mathbb{C}$. Then, as before, the joint empirical distribution of $\phi_j^{(n)}, j = 1, 2, \dots, n-1$, converges weakly to ν (which is the uniform law on $[0, 1]$), and so, the joint empirical distribution of $\exp(2\pi i \phi_j^{(n)}), j = 1, 2, \dots, n-1$, converges weakly to uniform on \mathcal{C} . Lemma 2.2 then gives us the desired result. \square

ACKNOWLEDGEMENTS

The author grateful to Robin Pemantle, Philip Gressman and Andreea Nicoara for stimulating discussions and helpful suggestions.

REFERENCES

- [Ch68] K. Chandrasekharan. *Introduction to Analytic Number Theory*. Springer-Verlag, 1968.
- [CN06] W. S. Cheung and T. W. Ng. A companion matrix approach to the study of zeros and critical points of a polynomial. *J. Math. Anal. Appl.*, 319(2): 690-707, 2006.
- [Co78] John B. Conway. *Functions of One Complex Variable*. Springer-Verlag, 1978.
- [dB46] N. G. de Bruijn. On the zeros of a polynomial and of its derivative. *Indag. Math.*, 8: 635-643, 1946.
- [dBS47] N. G. de Bruijn and T. A. Springer. On the zeros of a polynomial and of its derivative. II. *Indag. Math.*, 9: 264-270, 1947.
- [Pe12] R. Pemantle. Hyperbolicity and stable polynomials in combinatorics and probability. *Preprint: <http://www.math.upenn.edu/~pemantle/papers/Preprints/hyperbolic.pdf>*
- [PR12] R. Pemantle and I. Rivin. The distribution of the zeroes of the derivative of a random polynomial. *Preprint: <http://www.math.upenn.edu/~pemantle/papers/Preprints/zeros.pdf>*
- [RS02] Q. I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials*. Oxford University Press, Oxford, 2002.
- [St48] M. H. Stone. The Generalized Weierstrass Approximation Theorem. *Mathematics Magazine*, 21(4): 167-184, 1948.