THE COMPLEX ZEROS OF RANDOM POLYNOMIALS

LARRY A. SHEPP AND ROBERT J. VANDERBEI

ABSTRACT. Mark Kac gave an explicit formula for the expectation of the number, $\nu_n(\Omega)$, of zeros of a random polynomial,

$$P_n(z) = \sum_{j=0}^{n-1} \eta_j z^j,$$

in any measurable subset Ω of the *reals*. Here, $\eta_0, \ldots, \eta_{n-1}$ are independent standard normal random variables. In fact, for each $n > 1$, he obtained an explicit intensity function g_n for which

$$\mathbf{E}\nu_n(\Omega) = \int_\Omega g_n(x)dx.$$

Here, we extend this formula to obtain an explicit formula for the expected number of zeros in any measurable subset Ω of the complex plane $\mathbb{C}$. Namely, we show that

$$\mathbf{E}\nu_n(\Omega) = \int_\Omega h_n(x,y)dxdy + \int_{\Omega\cap\mathbb{R}} g_n(x)dx,$$

where h_n is an explicit intensity function. We also study the asymptotics of h_n showing that for large n its mass lies close to, and is uniformly distributed around, the unit circle.

1. INTRODUCTION.

More than fifty years ago, Mark Kac [7] gave an explicit formula for the expectation of the number, $\nu_n(\Omega)$, of zeros of a random polynomial,

$$P_n(z) = \sum_{j=0}^{n-1} \eta_j z^j, \quad z \in \mathbb{C}, \tag{1.1}$$

in any measurable subset Ω of the reals. Here, $\eta_0, \ldots, \eta_{n-1}$ are independent standard normal random variables (which Kac argues is the most natural choice for a "typical" polynomial since this distribution is invariant under orthogonal transformations - but of course other choices for the measure are also interesting). In fact, for each $n > 1$, he obtained an explicit intensity function g_n for which

$$\mathbf{E}\nu_n(\Omega) = \int_\Omega g_n(x)dx.$$

1991 *Mathematics Subject Classification.* Primary 30C15, Secondary 30B20, 26C10, 60B99.

Second author's research supported by AFOSR through grant AFOSR-91-0359.

Here, we extend this result by deriving an explicit formula for the expected number of zeros in any measurable subset Ω of the complex plane $\mathbb{C}$. Namely, we show that

$$\mathbf{E}\nu_n(\Omega) = \int_\Omega h_n(x,y)dxdy + \int_{\Omega\cap\mathbb{R}} g_n(x)dx,$$

where h_n is an explicit intensity function. (Note that we make the usual identification between a point z of the complex plane and its real and imaginary parts, x and y.)

The intensity function h_n is conveniently expressed in terms of the following three real-valued functions defined on $\mathbb{C}$

$$B_k(z) = \sum_{j=0}^{n-1} j^k |z|^{2j}, \quad z \in \mathbb{C}, \quad k = 0,1,2,$$

and the following two complex-valued functions

$$A_k(z) = \sum_{j=0}^{n-1} j^k z^{2j}, \quad z \in \mathbb{C}, \quad k = 0,1.$$

Finally, let

$$D_0(z) = \sqrt{B_0^2(z) - |A_0|^2(z)}. \tag{1.2}$$

Our main result is

Theorem 1.1. *For each region $\Omega \in \mathbb{C}$,*

$$\mathbf{E}\nu_n(\Omega) = \int_\Omega h_n(x,y)dxdy + \int_{\Omega\cap\mathbb{R}} g_n(x)dx, \tag{1.3}$$

where

$$h_n = \frac{B_2 D_0^2 - B_0(B_1^2 + |A_1|^2) + B_1(A_0\bar{A}_1 + \bar{A}_0 A_1)}{\pi |z|^2 D_0^3},$$

and

$$g_n = \frac{(B_0B_2 - B_1^2)^{1/2}}{\pi |z| B_0}.$$

As already mentioned, Kac [7] was the first to study the real zeros for random polynomials of this type and he obtained the intensity function g_n. Somewhat earlier (but delayed in publication by the war), Rice [12] obtained a similar formula for the expected number of zeros of a general stochastic process $X(t)$ again depending on a real parameter t but did not consider the special case when X is a polynomial in t. Here, we obtain Kac's g_n as a consequence of our analysis of the zeros in the complex plane.

While the definition of h_n in Theorem 1.1 looks rather complicated, it is nevertheless amenable both to computation (see Figures 1 to 6) and to asymptotic analysis. Indeed, the computed plots appearing on the left-hand side in Figures 1 to 6 clearly show that as n gets large the zeros tend to lie very close to the unit circle and, ignoring the real roots, appear to be approximately uniformly distributed around the circle. Regarding analysis, the next Theorem shows that the intensities h_n and g_n have well-defined limits as n tends to infinity. These limits are best expressed in terms of the following functions:

$$B(z) = \frac{1}{1-|z|^2}, \tag{1.4}$$

$$A(z) = \frac{1}{1-z^2}, \tag{1.5}$$

and

$$D(z) = \sqrt{B^2(z) - |A|^2(z)}. \tag{1.6}$$

Theorem 1.2. *We have the following limits:*

$$h = \lim_n h_n = \frac{|B|D}{\pi}$$

and

$$g = \lim_n g_n = \frac{1}{\pi}|B|.$$

The limit functions, h and g, can be viewed as intensity functions for the zeros of the random power series

$$P(z) = \sum_{j=0}^{\infty} \eta_j z^j$$

(at least within its radius of convergence, $|z| < 1$).

Following the initial work of Kac and Rice, a large body of research on zeros of random polynomials has appeared – see [2] for a fairly complete account including an extensive list of references. Most of this work has focused on the real zeros; [4], [5] and [14] being a few notable exceptions. Also, the recent paper of Edelman and Kostlan [3] gives a very elegent geometric treatment of the problem.

In the following section, we derive explicit formulas for the intensity functions h_n and g_n. Then in Section 3, we study the limiting form of these functions as n tends to infinity. Section 4 presents more delicate asymptotic analysis that shows how h_n and g_n behave when n is large, but finite. Next, in Section 5, we discuss computational issues and finally, in Section 6 we offer some speculation and suggest future research directions.

We end this section by remarking that even though we usually consider the entire complex plane, it suffices in general to study simply the unit disk centered at the origin since, given a random polynomial $P_n(z)$, the transformation to $z^n P_n(1/z)$ produces a new random polynomial whose zeros are exactly the reciprocals of the zeros of $P_n(z)$. Hence, for any set A, we have that

$$\mathbf{E}\nu_n(A^{-1}) = \mathbf{E}\nu_n(A), \tag{1.7}$$

where $A^{-1} = \{z : z^{-1} \in A\}$. From this it follows easily that any intensity function, say h must satisfy

$$h(z) = h(1/z)\frac{1}{|z|^2}. \tag{1.8}$$

2. The Intensity Functions h_n and g_n.

This section is devoted to the proof of Theorem 1.1. We begin with

Proposition 2.1. *For each region $\Omega \in \mathbb{C}$ whose boundary intersects the real axis at most only finitely many times,*

$$\mathbf{E}\nu_n(\Omega) = \frac{1}{2\pi i}\int_{\partial\Omega} \frac{1}{z} F(z) dz, \tag{2.1}$$

where

$$F = \frac{B_1 D_0 + B_0 B_1 - \bar{A}_0 A_1}{B_0 D_0 + B_0^2 - \bar{A}_0 A_0}. \tag{2.2}$$

Proof. The argument principle (see, e.g., [1], p. 151) gives an explicit formula for the random variable $\nu_n(\Omega)$, namely

$$\nu_n(\Omega) = \frac{1}{2\pi i}\int_{\partial\Omega} \frac{P_n'(z)}{P_n(z)} dz. \tag{2.3}$$

Taking expectations in (2.3) and then interchanging expectation and contour integration (the justification of which is tedious but doable), we get

$$\mathbf{E}\nu_n(\Omega) = \frac{1}{2\pi i}\int_{\partial\Omega} \mathbf{E}\frac{P_n'(z)}{P_n(z)} dz. \tag{2.4}$$

To facilitate computations, it is advantagous to multiply and divide by z (multiplying inside the expectation and dividing outside). The following Lemma shows that away from the real axis the function

$$F(z) = \mathbf{E}\frac{z P_n'(z)}{P_n(z)} \tag{2.5}$$

simplifies to the expression given in (2.2) and, since we've assumed that $\partial\Omega$ intersects the real axis at only finitely many points, this finishes the proof. $\square$

Lemma 2.2. *Let F denote the function defined by* (2.5). *Off the real axis,*

$$F = \frac{B_1 D_0 + B_0 B_1 - \bar{A}_0 A_1}{B_0 D_0 + B_0^2 - \bar{A}_0 A_0}.$$

On the real axis, F has a jump discontinuity. Indeed, for each $x \in \mathbb{R}$,

$$\lim_{z \to x:\ \mathrm{Im}(z)>0} F = \frac{B_1 - i(B_0 B_2 - B_1^2)^{1/2}}{B_0}$$

and

$$\lim_{z \to x:\ \mathrm{Im}(z)<0} F = \frac{B_1 + i(B_0 B_2 - B_1^2)^{1/2}}{B_0}.$$

Proof. First we consider the case where z does not lie on the real axis. Note that $P_n(z)$ and $zP_n'(z)$ are complex Gaussian random variables. It is convenient to work with their real and imaginary parts,

$$P_n(z) = \xi_1 + i\xi_2, \tag{2.6}$$

$$zP_n'(z) = \xi_3 + i\xi_4, \tag{2.7}$$

which are just linear combinations of the original standard normal random variables:

$$\begin{array}{rclrcl} \xi_1 & = & \sum_{j=0}^{n-1} a_j \eta_j, & \quad \xi_2 & = & \sum_{j=0}^{n-1} b_j \eta_j, \\ \xi_3 & = & \sum_{j=0}^{n-1} c_j \eta_j, & \quad \xi_4 & = & \sum_{j=0}^{n-1} d_j \eta_j. \end{array}$$

The coefficients in these linear combinations are given by

$$\begin{array}{rclcl} a_j & = & \mathrm{Re}(z^j) & = & \dfrac{z^j + \bar{z}^j}{2}, \\ b_j & = & \mathrm{Im}(z^j) & = & \dfrac{z^j - \bar{z}^j}{2i}, \\ c_j & = & j a_j, & & \\ d_j & = & j b_j. & & \end{array} \tag{2.8}$$

Put $\xi = [\ \xi_1 \quad \xi_2 \quad \xi_3 \quad \xi_4\ ]^T$. The covariance among these four Gaussian random variables is easy to compute:

$$\mathrm{Cov}(\xi) = \mathbf{E}\xi\xi^T = \begin{bmatrix} a^T a & a^T b & a^T c & a^T d \\ b^T a & b^T b & b^T c & b^T d \\ c^T a & c^T b & c^T c & c^T d \\ d^T a & d^T b & d^T c & d^T d \end{bmatrix} \tag{2.9}$$

We now represent these four correlated Gaussian random variables in terms of four standard normals. To this end, we seek a lower triangular matrix $L = [\ l_{ij}\ ]$ such that the vector ξ is equal in distribution to $L\zeta$, where $\zeta = [\ \zeta_1 \quad \zeta_2 \quad \zeta_3 \quad \zeta_4\ ]^T$ is a vector of four independent standard normal random variables. The following simple calculation shows that L is the Cholesky factor for the covariance matrix:

$$\mathrm{Cov}(\xi) = \mathbf{E}\xi\xi^T = \mathbf{E}L\zeta\zeta^T L^T = LL^T. \tag{2.10}$$

Now, since $\xi \stackrel{\mathrm{D}}{=} L\zeta$ and L is lower triangular (the symbol $\stackrel{\mathrm{D}}{=}$ denotes equality in distribution), we get that

$$\begin{aligned} \frac{zP_n'(z)}{P_n(z)} &= \frac{\xi_3 + i\xi_4}{\xi_1 + i\xi_2} \\ &\stackrel{\mathrm{D}}{=} \frac{(l_{31} + il_{41})\zeta_1 + (l_{32} + il_{42})\zeta_2 + (l_{33} + il_{43})\zeta_3 + il_{44}\zeta_4}{(l_{11} + il_{21})\zeta_1 + il_{22}\zeta_2}. \end{aligned}$$

Hence, exploiting the independence of the ζ_i's, we see that

$$F(z) = \mathbf{E}\frac{zP_n'(z)}{P_n(z)} = \mathbf{E}\frac{\alpha\zeta_1 + \beta\zeta_2}{\gamma\zeta_1 + \delta\zeta_2}, \tag{2.11}$$

where

$$\begin{aligned} \alpha &= l_{31} + il_{41} \qquad & \beta &= l_{32} + il_{42} \\ \gamma &= l_{11} + il_{21} \qquad & \delta &= il_{22}. \end{aligned}$$

Splitting up the numerator in (2.11) and exploiting the exchangeability of ζ_1 and ζ_2, we can rewrite the expectation as follows:

$$F(z) = \frac{\alpha}{\delta} f(\gamma/\delta) + \frac{\beta}{\gamma} f(\delta/\gamma),$$

where f is a complex-valued function defined on $\mathbb{C} \setminus \mathbb{R}$ by

$$f(w) = \mathbf{E}\frac{\zeta_1}{w\zeta_1 + \zeta_2}.$$

The expectation appearing in the definition of f can be explicitly computed. Indeed,

$$\begin{aligned} f(w) &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} \frac{\rho\cos\theta}{w\rho\cos\theta + \rho\sin\theta} e^{-\rho^2/2}\rho d\rho d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{d\theta}{w + \tan\theta} \end{aligned}$$

and the substitution $u = \tan\theta$ can then be used to rewrite this integral as

$$f(w) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{(w+u)(1+u^2)} du$$

which is easily solved using the calculus of residues. The final result is that

$$f(w) = \begin{cases} \dfrac{1}{w+i}, & \mathrm{Im}(w) > 0, \\ \dfrac{1}{w-i}, & \mathrm{Im}(w) < 0. \end{cases}$$

Recalling the definition of δ and γ, we see that

$$\frac{\gamma}{\delta} = \frac{l_{21}}{l_{22}} - i\frac{l_{11}}{l_{22}}.$$

In general, l_{11} and l_{22} are just nonnegative. However, it is not hard to show that they are both strictly positive whenever z has a nonzero imaginary part. Hence, γ/δ lies in the lower half-plane, δ/γ lies in the upper half-plane, and

$$\begin{aligned} F(z) &= \frac{\alpha}{\delta}\frac{1}{\frac{\gamma}{\delta} - i} + \frac{\beta}{\gamma}\frac{1}{\frac{\delta}{\gamma} + i} \\ &= \frac{i\alpha + \beta}{i\gamma + \delta} \\ &= \frac{l_{32} - l_{41} + i(l_{31} + l_{42})}{-l_{21} + i(l_{11} + l_{22})}. \end{aligned} \tag{2.12}$$

At this point, we need explicit formulas for the elements of the Cholesky factor L. From (2.9) and (2.10), we see that

$$\begin{aligned} a^T a &= l_{11}^2 \\ b^T a &= l_{21} l_{11} \\ c^T a &= l_{31} l_{11} \\ d^T a &= l_{41} l_{11} \\ b^T b &= l_{21}^2 + l_{22}^2 \\ c^T b &= l_{31} l_{21} + l_{32} l_{22} \\ d^T b &= l_{41} l_{21} + l_{42} l_{22}. \end{aligned}$$

Solving these equations in succession, we get

$$\begin{aligned}
l_{11} &= \frac{a^T a}{\sqrt{a^T a}} \\
l_{21} &= \frac{b^T a}{\sqrt{a^T a}} \\
l_{31} &= \frac{c^T a}{\sqrt{a^T a}} \\
l_{41} &= \frac{d^T a}{\sqrt{a^T a}} \\
l_{22} &= \frac{(a^T a)(b^T b) - (b^T a)^2}{\sqrt{a^T a} R} \\
l_{32} &= \frac{(a^T a)(c^T b) - (c^T a)(b^T a)}{\sqrt{a^T a} R} \\
l_{42} &= \frac{(a^T a)(d^T b) - (d^T a)(b^T a)}{\sqrt{a^T a} R},
\end{aligned}$$

where

$$R = \sqrt{(a^T a)(b^T b) - (b^T a)^2}.$$

Substituting these expressions into (2.12) and simplifying, we see that

$$F(z) = \frac{-d^T a + i c^T a - i \left(a^T a(-d^T b + i c^T b) - (-d^T a + i c^T a) b^T a \right) / R}{-b^T a + i a^T a + i R}. \tag{2.13}$$

Recalling the definitions of a_j, b_j, c_j, and d_j given in (2.8), it is easy to check that the following identities hold:

$$\begin{aligned}
a^T a &= +\frac{1}{4}(A_0 + 2B_0 + \bar{A}_0), \\
b^T a &= -\frac{i}{4}(A_0 - \bar{A}_0), \\
c^T a &= +\frac{1}{4}(A_1 + 2B_1 + \bar{A}_1), \\
d^T a &= -\frac{i}{4}(A_1 - \bar{A}_1), \\
b^T b &= -\frac{1}{4}(A_0 - 2B_0 + \bar{A}_0), \\
c^T b &= -\frac{i}{4}(A_1 - \bar{A}_1), \\
d^T b &= -\frac{1}{4}(A_1 - 2B_1 + \bar{A}_1).
\end{aligned}$$

Plugging these expressions into (2.13) and simplifying, we get that

$$F(z) = \frac{A_1 + B_1 + (A_0 B_1 + B_0 B_1 - A_1 B_0 - \bar{A}_0 A_1)/D_0}{A_0 + B_0 + D_0}, \tag{2.14}$$

where D_0 is given in (1.2). It turns out that further simplification occurs if we make the denominator real by the usual technique of multiplying and dividing by its complex conjugate. We leave out the algebraic details except to mention that a factor of $A_0 + 2B_0 + \bar{A}_0$ cancels out from the numerator and denominator leaving us with

$$F(z) = \frac{B_1 D_0 + B_0 B_1 - \bar{A}_0 A_1}{D_0(B_0 + D_0)}, \tag{2.15}$$

or, expanding out D_0^2,

$$F(z) = \frac{B_1 D_0 + B_0 B_1 - \bar{A}_0 A_1}{B_0 D_0 + B_0^2 - \bar{A}_0 A_0}. \tag{2.16}$$

The expression in (2.15) is as simple as possible, whereas (2.16) illustrates a certain parallelism between the numerator and denominator.

Now consider a point x on the real axis. On the reals, $A_k = B_k$, $k = 0, 1$, and so $D_0 = 0$. Hence, the right-hand side in (2.15) is an indeterminate form. To analyze the limiting behavior of F near the real axis, we first divide the numerator and denominator by D_0:

$$F = \frac{B_1 + \dfrac{B_0 B_1 - \bar{A}_0 A_1}{D_0}}{B_0 + D_0}. \tag{2.17}$$

Now, only the ratio in the numerator is indeterminate. To study it, we write $z = re^{i\theta}$ and investigate the numerator and denominator of this ratio when θ is small. From the definitions, we see that

$$\begin{aligned} B_0 B_1 - \bar{A}_0 A_1 &= \sum_{j,k} k r^{2(k+j)} (1 - e^{2(k-j)\theta i}) \\ &= \sum_{j,k} k r^{2(k+j)} 2(j-k)\theta i + o(\theta) \\ &= 2\theta i (B_1^2 - B_0 B_2) + o(\theta) \end{aligned} \tag{2.18}$$

and

$$\begin{aligned} B_0^2 - |A_0|^2 &= \sum_{j,k} r^{2(k+j)} (1 - e^{2(k-j)\theta i}) \\ &= -\sum_{j,k} r^{2(k+j)} 2(k-j)^2 \theta^2 i^2 + o(\theta^2) \\ &= 4\theta^2 (B_0 B_2 - B_1^2) + o(\theta^2). \end{aligned} \tag{2.19}$$

Hence, recalling that $D_0 = \sqrt{B_0^2 - |A_0|^2}$, we see that

$$\begin{aligned} \frac{B_0 B_1 - \bar{A}_0 A_1}{D_0} &= \operatorname{sgn}(\theta) i \frac{B_1^2 - B_0 B_2 + o(\theta)}{(B_0 B_2 - B_1^2 + o(\theta^2))^{1/2}} \\ &= -\operatorname{sgn}(\theta) i (B_0 B_2 - B_1^2)^{1/2} + o(\theta). \end{aligned} \tag{2.20}$$

Combining (2.17) and (2.20), we get the desired limits expressing the jump discontinuity on the real axis. □

Proof of Theorem 1.1. Without loss of generality, it suffices to consider regions Ω that are either regions that do not intersect the real axis or polar rectangles that do intersect the real axis. We begin by considering a region Ω that does not intersect the real axis. Applying Stokes' theorem to the expression for $\mathbf{E}\nu_n(\Omega)$ given in Proposition 2.1, we see that

$$\mathbf{E}\nu_n(\Omega) = \frac{1}{\pi} \int_\Omega \frac{1}{z} \frac{\partial}{\partial \bar{z}} F(z, \bar{z}) dx dy.$$

Note that we are now writing $F(z, \bar{z})$ to emphasize the fact that F depends on both z and $\bar{z}$. Letting the dagger symbol stand for the derivative with respect to $\bar{z}$, we see from Lemma 2.2 that

$$\begin{aligned} \frac{\partial F}{\partial \bar{z}} = & \Big\{ (B_0 D_0 + B_0^2 - \bar{A}_0 A_0)(B_1^\dagger D_0 + B_1 D_0^\dagger + B_0^\dagger + B_0 B_1^\dagger - \bar{A}_0^\dagger A_1) \\ & -(B_1 D_0 + B_0 B_1 - \bar{A}_0 A_1)(B_0^\dagger D_0 + B_0 D_0^\dagger + 2 B_0 B_0^\dagger - \bar{A}_0^\dagger A_0) \Big\} \\ & /(B_0 D_0 + B_0^2 - |A_0|^2)^2. \end{aligned}$$

From their definitions, it is easy to verify that

$$B_0^\dagger = \frac{B_1}{\bar{z}}, \quad \bar{A}_0^\dagger = \frac{2\bar{A}_1}{\bar{z}}, \quad B_1^\dagger = \frac{B_2}{\bar{z}}.$$

Recalling that $D_0 = \sqrt{B_0^2 - |A|^2}$, we get that

$$D_0^\dagger = \frac{B_0 B_1 - A_0 \bar{A}_1}{\bar{z} D_0}.$$

Substituting these formulas for the derivatives into the expression given above for $\partial F / \partial \bar{z}$ and then engaging in tedious algebraic simplifications (a computer algebra package such as Mathematica would be useful here, but we confess that we did this by hand), we eventually arrive at the fact that $(\pi z)^{-1} \partial F(z, \bar{z}) / \partial \bar{z}$ equals the expression given in Theorem 1.1 for h_n.

Now consider a polar rectangle that covers a portion of the real axis. In other words, let Ω be the angular interval $(-\theta, \theta)$ crossed with a radial interval (r_0, r_1) (we have assumed without loss of generality that the polar rectangle does not intersect the negative part of the real axis). Writing down the contour integral for $\mathbf{E}\nu_n(\Omega)$ given by Proposition 2.1 and letting θ tend to 0, we see that

$$\mathbf{E}\nu_n((r_0, r_1)) = \frac{1}{2\pi i}\int_{r_0}^{r_1} \frac{F(r-) - F(r+)}{r} dr,$$

where $\nu_n((r_0, r_1))$ denotes the number of zeros in the interval (r_0, r_1) of the real axis and

$$F(r-) = \lim_{z \to r : \mathrm{Im}(z) < 0} F(z) \quad \text{and} \quad F(r+) = \lim_{z \to r : \mathrm{Im}(z) > 0} F(z).$$

Hence, from Lemma 2.2, we see that

$$g_n(r) = \frac{1}{2\pi i}\frac{F(r-) - F(r+)}{r} = \frac{\sqrt{B_0 B_2 - B_1^2}}{\pi r B_0}.$$

This completes the proof. □

3. LIMITING INTENSITY.

In this section we study the limits obtained by letting n tend to infinity. As mentioned in the introduction, these limits can be viewed as intensity functions for the limiting random power series. We begin by proving Theorem 1.2:

Proof of Theorem 1.2. The sums used to define the functions B_k and A_k can be summed explicitly. Recalling the functions A and B defined by (1.5) and (1.4), respectively, we have

$$B_0 = B - |z|^{2n} B, \tag{3.1}$$

$$B_1 = |z|^2 B^2 - \left(n|z|^{2n} B + |z|^{2n+2} B^2\right), \tag{3.2}$$

$$B_2 = |z|^2(1 + |z|^2) B^3 - \left(n^2 |z|^{2n} B + 2n|z|^{2n+2} B^2 + |z|^{2n+2}(1 + |z|^2) B^3\right), \tag{3.3}$$

$$A_0 = A - z^{2n} A,$$

$$A_1 = z^2 A^2 - \left(n z^{2n} A + z^{2n+2} A^2\right).$$

For $|z| < 1$, we have the following limits:

$$\lim_{n\to\infty} B_0 = B, \tag{3.4}$$

$$\lim_{n\to\infty} B_1 = |z|^2 B^2, \tag{3.5}$$

$$\lim_{n\to\infty} B_2 = |z|^2(1 + |z|^2) B^3, \tag{3.6}$$

$$\lim_{n\to\infty} A_0 = A,$$

$$\lim_{n\to\infty} A_1 = z^2 A^2.$$

Using these limits, it is easy to check that

$$\lim_{n\to\infty} \left(B_2 B_0^2 - B_0 B_1^2\right) = |z|^2 B^5,$$

and

$$\lim_{n\to\infty} \left(-B_2|A_0|^2 - B_0|A_1|^2 + B_1(A_0\bar{A}_1 + \bar{A}_0 A_1)\right) \\ = -|z|^2 B|A|^2 \left(B^2 + |\bar{z}B - zA|^2\right).$$

Then, using the definitions of A, B, and D, we see that

$$|\bar{z}B - zA|^2 = D^2.$$

Substituting these ingredients into the expression given in Theorem 1.1 for h_n and simplifying, we find that

$$\begin{aligned} \lim_{n\to\infty} h_n &= \frac{|z|^2 B^5 - |z|^2 B|A|^2 \left(B^2 + |\bar{z}B - zA|^2\right)}{\pi|z|^2 D^3} \\ &= \frac{BD}{\pi}. \end{aligned} \tag{3.7}$$

The case $|z| > 1$ is more tedious (and less interesting!). We can either refer to the inversion symmetry formula (1.8) or we can compute in a manner similar to the $|z| < 1$ case. Opting for the detailed computation, we note that this time we must keep track of terms involving the three highest orders: $n^2|z|^{2n}$, $n|z|^{2n}$, and $|z|^{2n}$. Then we compute the numerator of the right-hand factor in the expression given in Theorem 1.1 for h_n and we discover that the two highest-order terms cancel out leaving terms of the order $|z|^{2n}$ as the highest remaining order. Indeed,

$$B_2 B_0^2 - B_0 B_1^2 = -|z|^{6n}|z|^2 B^5 + o(|z|^{6n}),$$

$$-B_2|A_0|^2 - B_0|A_1|^2 + B_1(A_0\bar{A}_1 + \bar{A}_0 A_1) = \\ |z|^{6n}|z|^2 B|A|^2 \left(B^2 + |\bar{z}B - zA|^2\right) + o(|z|^{6n}),$$

and

$$D_0^2 = |z|^{4n} D^2 + o(|z|^{4n}).$$

Hence, the numerator and denominator of the right-hand factor both have $|z|^{6n}$ as their highest order terms, so this factor can be canceled. For the left-hand factor, both numerator and denominator have $|z|^{4n}$ as their highest order terms. Eliminating these and passing to the limit, we get

$$\lim_{n\to\infty} h_n = \frac{-BD}{\pi}. \tag{3.8}$$

Bearing in mind that $B > 0$ for $|z| < 1$ and $B < 0$ for $|z| > 1$, we see that (3.7) and (3.8) can be combined to give the formula asserted for h.

The same method is used to find $\lim_{n\to\infty} g_n$. Indeed, for $|z| < 1$, we use (3.4), (3.5), and (3.6) to see that

$$\lim_{n\to\infty} g_n = \frac{1}{\pi} B.$$

For $|z| > 1$, we retain the three highest order terms from (3.1), (3.2), and (3.3). As before the two highest order terms cancel out in the expression for $B_0 B_2 - B_1^2$ leaving us with

$$B_0 B_2 - B_1^2 = |z|^{4n} |z|^2 B^4 + o(|z|^{4n}).$$

From this expression, we easily see that

$$\lim_{n\to\infty} g_n = -\frac{1}{\pi} B,$$

which completes the proof. □

4. Asymptotics for the Number of Roots.

In this section we derive limiting expressions for the expected number of zeros in disks and sectors. We start with the disk, $B(r)$, of radius r centered at 0.

Theorem 4.1. *For* $r < 1$,

$$\lim_{n\to\infty} \mathbf{E}\nu_n(B(r)) = \frac{r^2}{1-r^2} + O(1).$$

Proof. From Proposition 2.1, we have that

$$\begin{aligned} \mathbf{E}\nu_n(B(r)) &= \frac{1}{2\pi i} \int_{\partial B(r)} \frac{1}{z} F(z) dz \\ &= \frac{1}{2\pi} \int_0^{2\pi} F(z) d\theta. \end{aligned}$$

Now, applying the bounded convergence theorem, we see that

$$\lim_{n\to\infty} \mathbf{E}\nu_n(B(r)) = \frac{1}{2\pi} \int_0^{2\pi} \lim_{n\to\infty} F(z) d\theta$$

and, using the expression given in (2.15) for F, we get that

$$\lim_{n\to\infty} F(z) = \frac{|z|^2 B^2 D + B|z|^2 B^2 - \bar{A} z^2 A^2}{(B+D)D}.$$

Adding and subtracting $|z|^2 B |A|^2$ from the numerator and using the fact that $|z|^2 B = r^2/(1-r^2)$ on the contour of integration, we see that

$$\lim_{n\to\infty} F(z) = \frac{r^2}{1-r^2} + \frac{|A|^2}{(B+D)D} \left(|z|^2 B - z^2 A \right). \tag{4.1}$$

From the definitions of A, B, and D, it is easy to check that

$$|z|^2 B - z^2 A = \left(|z|^2 - z^2\right) AB$$

and that

$$D = 2|\mathrm{Im}(z)|B|A|.$$

Hence, the second term on the right in (4.1) simplifies to

$$(1-r^2)\frac{|A|A(r^2-z^2)}{2\left(1+2|\mathrm{Im}(z)||A|\right)|\mathrm{Im}(z)|} \tag{4.2}$$

and so the proof will be complete if we show that the integral of this expression over the circle of radius r is bounded. To this end, we note that the triangle inequality combined with the fact that $|\sin\theta| = |1-e^{2i\theta}|/2$ gives us

$$\left|\int_0^{2\pi} \frac{|A|A(r^2-z^2)}{2\left(1+2|\mathrm{Im}(z)||A|\right)|\mathrm{Im}(z)|} d\theta\right| \le r\int_0^{2\pi} \frac{1}{|1-r^2e^{2i\theta}|^2} d\theta \tag{4.3}$$

The integral on the right-hand side can be integrated explicitly:

$$\int_0^{2\pi} \frac{1}{|1-r^2e^{2i\theta}|^2} d\theta = \frac{2\pi}{(1-r^2)(1+r^2)}. \tag{4.4}$$

Combining (4.4), (4.3), and (4.2), we get that the second term in (4.1) is bounded by $2\pi r/(1+r^2)$. This completes the proof. □

Theorem 4.2. *For $s \ge 0$,*

$$\lim_{n\to\infty} \mathbf{E}\frac{1}{n}\nu_n\left(B(e^{-s/2n})\right) = \frac{(1-e^{-s}(1+s))}{s(1-e^{-s})} = \frac{1}{2} - \frac{s}{12} + o(s).$$

Proof. We begin as we began the proof of the previous theorem with

$$\mathbf{E}\nu_n(B(r)) = \frac{1}{2\pi}\int_0^{2\pi} F(z)d\theta,$$

where

$$F(z) = \frac{B_1 + \frac{B_0B_1 - \bar{A}_0A_1}{D_0}}{B_0 + D_0},$$

and this time

$$r = e^{-s/2n},$$

and

$$z = re^{i\theta}.$$

The theorem then follows from the dominated convergence theorem and the following easy-to-derive asymptotic formulas (the last two holding everywhere except $\theta = 0$ and $\theta = \pi$):

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} B_0 = \lim_{n\to\infty} \frac{1}{n} D_0 &= (1-e^{-s})/s,\\ \lim_{n\to\infty} \frac{1}{n^2} B_1 &= \left(1-e^{-s}(1+s)\right)/s^2,\\ \lim_{n\to\infty} \frac{1}{n} A_0 = \lim_{n\to\infty} \frac{1}{n} A_1 &= 0.\end{aligned}$$

□

Let $S(\theta_0, \theta_1)$ denote the sector of the complex plane consisting of all points whose argument lies between θ_0 and θ_1. The next theorem shows that asymptotically the zeros are uniformly distributed in argument. Of course, the $2/\pi \log n$ real zeros disappear since we are normalizing by dividing by n.

Theorem 4.3. *For each sector $S(\theta_0, \theta_1)$ that does not intersect the real axis,*

$$\lim_{n\to\infty} \frac{1}{n} \mathbf{E}\nu_n(S(\theta_0,\theta_1)) = \frac{\theta_1 - \theta_0}{2\pi}.$$

Šparo and Šur [] have obtained the analogous result for complex Gaussians using a very general and elegent result of Erdös and Turán [] (see also [], Theorem 8.1).

Proof. According to (1.7), it suffices to prove the analogous limit for the intersection, $P(\theta_0, \theta_1)$, of $S(\theta_0, \theta_1)$ with the unit disk. By Proposition 2.1,

$$\mathbf{E}\nu_n(P(\theta_0,\theta_1)) = \frac{1}{2\pi i} \int_0^1 \frac{1}{r} \left(F(re^{i\theta_0}) - F(re^{i\theta_1})\right) dr + \frac{1}{2\pi} \int_{\theta_0}^{\theta_1} F(e^{i\theta}) d\theta,$$

where F is given by (2.2). Note that the first integrand is bounded uniformly in n even at $r = 0$. To see this, we take note of the following limits:

$$\begin{gathered}\lim_{r\to 0} B_0 = \lim_{r\to 0} A_0 = 1,\\ \lim_{r\to 0} D_0 = 0,\\ \lim_{r\to 0} \frac{1}{r} B_1 = 1,\end{gathered}$$

and

$$\lim_{r\to 0} \frac{1}{r} \frac{B_0 B_1 - \bar{A}_0 A_1}{D_0} = -\frac{e^{i\theta}}{2\sin\theta}.$$

This last limit is easy to obtain from expressions (2.18) and (2.19). Hence, from (2.17) we see that

$$\lim_{r\to 0} \frac{1}{r} F(re^{i\theta}) = 1 - \frac{e^{i\theta}}{2\sin\theta}$$

and so we can interchange limit and integration to see that in the limit $1/n$ times the first integral vanishes.

To study the limiting behavior of $1/n$ times the second integral, we first rewrite F as

$$F = \frac{B_1}{D_0} - \frac{\bar{A}_0 A_1}{D_0(B_0 + D_0)}.$$

Now, for $z = e^{i\theta}$,

$$B_0 = n, \quad B_1 = \frac{n(n-1)}{2}, \quad \text{and} \quad D_0 = \sqrt{n^2 - |A_0|^2}.$$

Since A_0 and A_1 are bounded on $P(\theta_0, \theta_1)$, it follows that

$$\lim_{n\to\infty} \frac{1}{n} \frac{B_1}{D_0} = \frac{1}{2}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \frac{\bar{A}_0 A_1}{D_0(B_0 + D_0)} = 0.$$

Therefore,

$$\lim_{n\to\infty} \frac{1}{n} F(e^{i\theta}) = \frac{1}{2}$$

and so the dominated convergence theorem implies that

$$\lim_{n\to\infty} \frac{1}{n} \mathbf{E}\nu_n(P(\theta_0, \theta_1)) = \frac{\theta_1 - \theta_0}{4\pi}.$$

This completes the proof. □

5. Numerical Computation.

In this section, we describe the computer program used to produce Figure 1 through Figure 5. Each figure shows two plots. On the left is a grey-scale plot of the intensity functions g_n and h_n. On the right is a plot of $20,000$ zeros obtained by generating random polynomials and explicitly finding their zeros.

The intensity plots appearing on the left were produced by partitioning the given square domain into a 256 by 256 grid of "pixels" (i.e., small squares) and computing the intensity function in the center of each pixel. The grey-scale was computed by assigning white to the pixel with the smallest value and black to the pixel with the largest value and then linearly interpolating all values in between. This grey-scale computation was performed separately for h_n and for g_n (which appears only on the x-axis) and so no conclusions should be drawn comparing the intensity shown on the x-axis with that shown off from it.

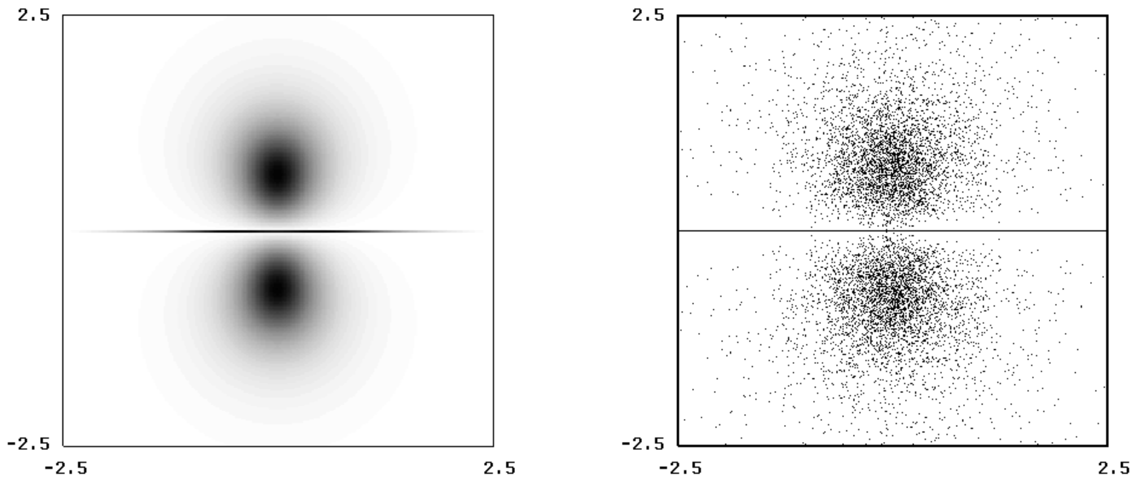

FIGURE 1. *Random quadratics* ($n = 3$). In this figure and the next several, the left-hand plot is a grey-scale image of the intensity functions h_n and g_n (which is concentrated on the x-axis). The right-hand plot shows $20,000$ roots from randomly generated polynomials. Subsequent captions give more information pertaining to each of these figures.

Of course, the intensity function g_n is one-dimensional and therefore it would be natural (and more informative) to make separate plots of values of g_n verses x, but such plots appear in many places (see, e.g., [8]) and so it seemed unnecessary to produce them here. Suffice it to say that g_n is a density whose mass lies primarily near ± 1.

The $20,000$ zeros plotted on the right were produced by a novel root finding algorithm, which we shall describe briefly. Given a polynomial P, the algorithm that finds its zeros in a given square R is recursive. Let S denote the smallest disk that contains R. Using the argument principle, we can compute the number of zeros in S:

$$\nu(S) = \frac{1}{2\pi i} \int_{\partial S} \frac{P'(z)}{P(z)} dz. \tag{5.1}$$

If $\nu(S)$ is zero, then we stop since there are then no zeros in R. Otherwise, we compute the sum of the zeros using the following analogue of the argument principle:

$$\sigma(S) = \sum_{i: z_i \in S} z_i = \frac{1}{2\pi i} \int_{\partial S} \frac{z P'(z)}{P(z)} dz. \tag{5.2}$$

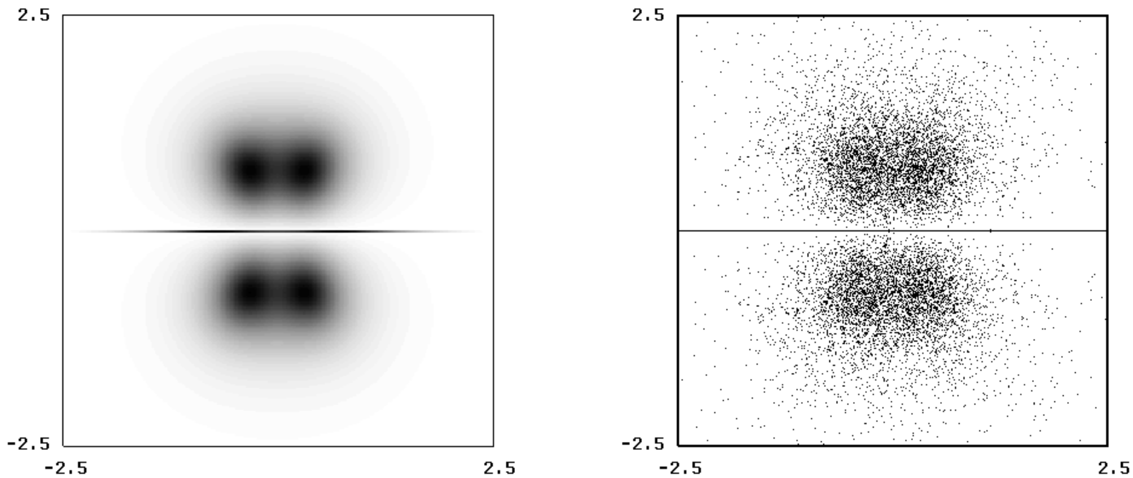

FIGURE 2. *Random cubics* ($n = 4$). Note that, for the left-hand plots, the grey-scales for h_n and g_n are produced separately and in such a way that both use the full range from white to black.

Here, the z_i's denote the zeros of P. The ratio

$$\alpha(S) = \sigma(S)/\nu(S)$$

is then the barycenter of the zeros contained in S (counting multiplicities). Next, let B be a disk of very small radius centered at $\alpha(S)$ and again using the argument principle we count the number of zeros in this disk. If this number equals $\nu(S)$, then there is only one zero in S, it is at $\alpha(S)$ and it has multiplicity $\nu(S)$. Hence, again we stop (but first, we check whether the zero is in fact in R and if it is, we add it to our list of zeros found so far). In all other cases, there are still two or more zeros to be found and so we partition the square into four subsquares and recursively call this same procedure on those subsquares. Except for discussing how to compute the contour integrals, this is a complete description of the algorithm used to produce the right-hand plots.

The contour integrals in (5.1) and (5.2) are computed as follows. First we compute $\nu(S)$. To do this, we begin by simply using a discrete approximation to the integral consisting of 8 points spread equally around the boundary of S. Then we check to see if $\nu(S)$ is close to being an integer. If it is, we are done. If it isn't, then we spread 8 more points midway between each point already used in the sum and we update the sum with these points. Again we check nearness to integrality. We continue in this way until the integrality condition is satisfied. The contour integral in (5.2) is then computed using the same number of points as was used in (5.1).

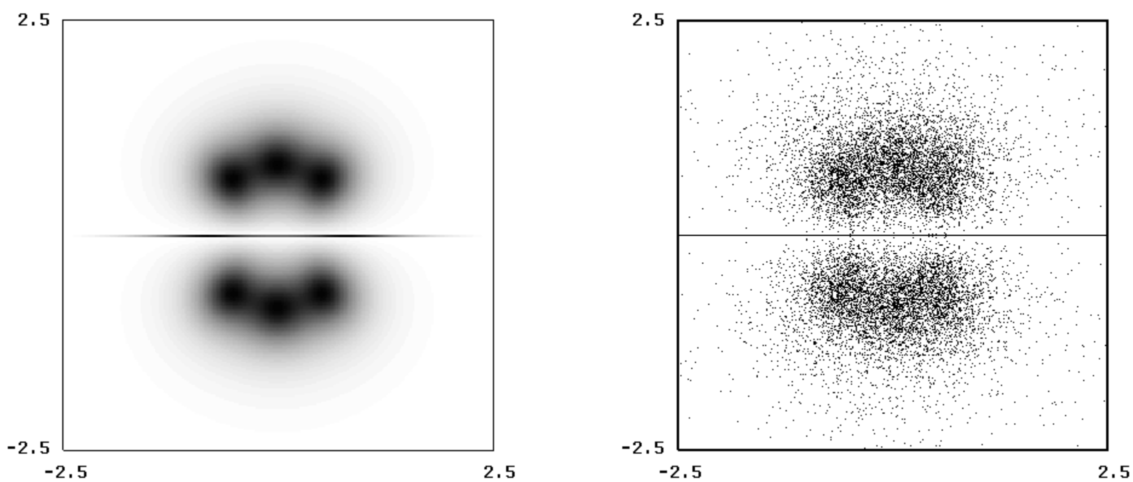

FIGURE 3. *Random quartics* ($n = 5$). Note that, for the right-hand plots, most if not all of the "pixels" on the x-axis have been hit by at least one root. A more accurate image for the x-axis would have been obtained had we used a grey-scale to indicate how many times each pixel on the x-axis was hit.

6. FINAL COMMENTS AND SUGGESTED FUTURE RESEARCH.

The machinery developed in this paper can be applied in many situations that we have not covered. For example, if the coefficients are assumed to be independent complex Gaussains (instead of real), then we can apply the same methods and in fact all computations are simpler. In this case, the intensity function does not have mass concentrated on the real axis (i.e., $g_n = 0$) and the intensity function is rotationally invariant. Furthermore, by going through the calculations, one discovers that h_n is just the square of Kac's intensity (with a different lead constant):

$$h_n = \frac{B_0 B_2 - B_1^2}{\pi |z|^2 B_0^2}.$$

The classic paper of Hammersley [5] covers this case (among others). Also, the random power series version was considered in [3].

In the case of real polynomials, Kac's formula has been extended to other distributions besides the normal law for the coefficients as well as in many other directions, see [2]. In view of the interest in zeros indicated by the enormous literature on the real zeros it is amazing that the simple results given here for the complex zeros have not been found earlier. One extension [10, 11] considers replacing the normal distribution by symmetric stable distributions, $S(\alpha)$, $0 < \alpha < 2$,

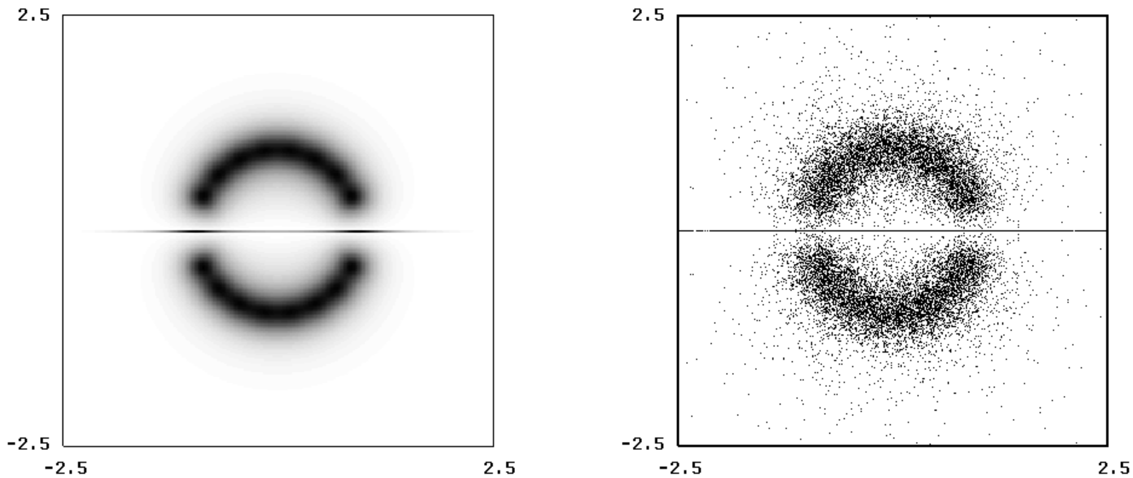

FIGURE 4. $n = 10$.

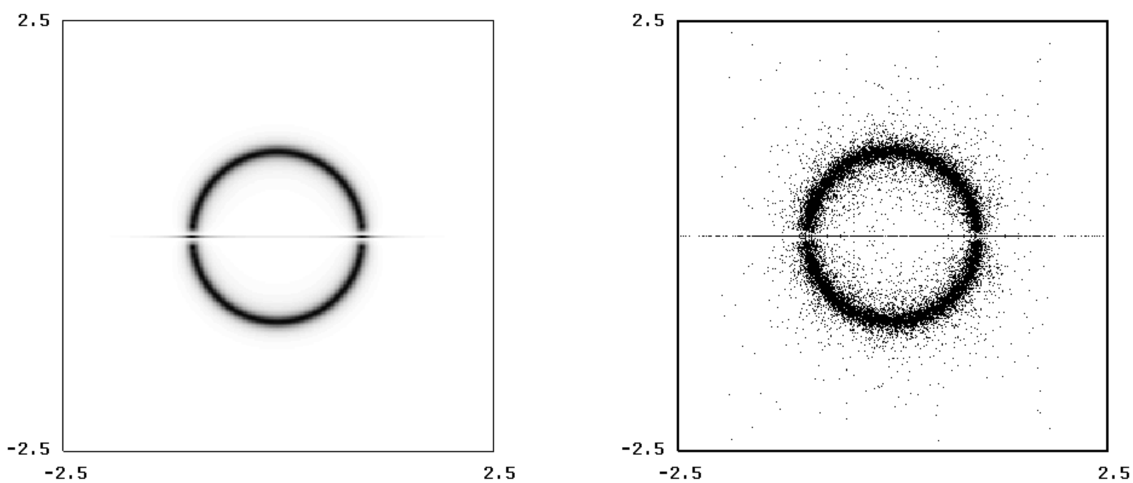

FIGURE 5. $n = 36$.

and obtains that the expected number of real zeros is asymptotic to $c(\alpha) \log n$ as n tends to infinity. The coefficient $c(\alpha)$ decreases from 1 to $2/\pi$ as α goes from 0 to 2, and is given by an explicit

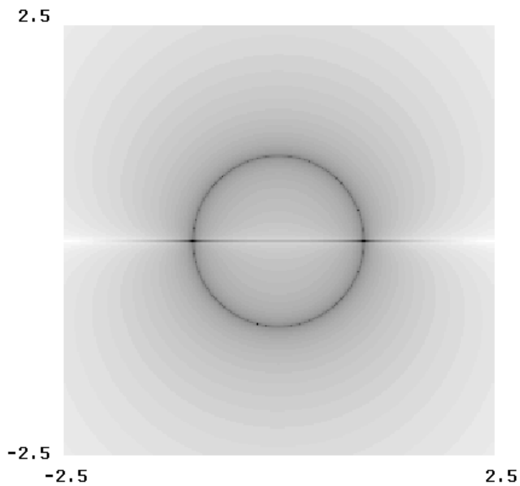

FIGURE 6. *Limiting intensity* (using a logarithmic grey-scale).

formula. Stevens [13] showed that if the coefficients are independent and are merely in the domain of attraction of the normal law, then the expected number of real zeros follows the same asymptotic, $\frac{2}{\pi}\log n$, and the corresponding result for the case when the coefficients are independent and in the domain of attraction of the stable law with parameter $1 < \alpha < 2$ was proved by Ibragimov and Maslova [6].

For the complex case, we do not see how to extend any of our results to domain of attraction statements. Moreover, we do not know what the change will be for $h_n(\alpha, x)$ if the coefficients are symmetric stable instead of normal. The techniques of [10, 11] do not seem to extend easily to the complex case. Needless to say, the same holds for coefficients merely within the domain of attraction of a symmetric stable law. However, we conjecture that the domain of attraction cases follow the same asymptotics as the corresponding stable law, in particular if $\alpha = 2$.

Let ρ_j denote the norm of the zero of $P_n(z)$ that is inside the unit circle and is the j-th closest to the unit circle. The results of this paper indicate that $\rho_j = 1 - X_j/n^2$, where the random variables $X_1, X_2, \ldots$ form a standard Poisson process. Similarly, the Kac formula for the real zeros suggests a similar conjecture except with a division by n instead of n^2. If these two statements are correct, then the complex zeros are much closer to the unit circle than the real zeros are to ± 1.

Assuming that the above conjectures are correct, there is a speculative corollary/application to Littlewood's problem of studying the behavior as n tends to infinity of the random variable defined as

$$m_n = \inf_{0 \le \theta \le 2\pi} \left| P_n(e^{i\theta}) \right|, \tag{6.1}$$

i.e., the infimum on the unit circle of $P_n(z)$. Here, the coefficients are assumed to be independent symmetric ± 1 random variables. Konyagin gave a new estimate for m_n in [9] by a rather complicated argument. We cannot obtain his delicate estimate by our methods, but there is at least some hope that this can be done provided that some further developments can be carried out. One would have to prove that the conjecture that the nearest complex zero to the unit circle is within $O(1/n^2)$ holds not only for normal coefficients but for +-1 coefficients. As pointed out privately by A.M. Odlyzko, it would then follow from Bernstein's theorem that an alternate proof to Konyagin's estimates in [9] would be obtainable.

Acknowledgement 6.1. The results of this paper and our resulting speculations on finding an alternate proof based on studying the zeros of $P_n(z)$ near the unit circle were stimulated by a talk by S.V. Konyagin. In the course of these speculations, the study of the zeros themselves began to be even more interesting to us that the application to Littlewood's problem. The application will have to await further inspiration and another paper (possibly by other authors).

We would also like to acknowledge Ofer Zeitouni for carefully reading an early draft of the paper and suggesting some important algebraic simplifications.

REFERENCES

1. L.V. Ahlfors, *Complex analysis*, McGraw-Hill, New York, NY, 1966.
2. A.T. Bharucha-Reid and M. Sambandham, *Random polynomials*, Academic Press, 1986.
3. A. Edelman and E. Kostlan, *How many zeros of a random polynomial are real?*, J. AMS (1995), To appear.
4. P. Erdös and P. Turán, *On the distribution of roots of polynomials*, Ann. Math. **51** (1950), 105–119.
5. J. Hammersley, *The zeros of a random polynomial*, Proc. Third Berkeley Symp. Math. Stat. Probability, vol. 2, 1956, pp. 89–111.
6. I.A. Ibragimov and N.B. Maslova, *On the expected number of real zeros of random polynomials. i. coefficients with zero means*, Teor. Veroyatn. Ee Primen. **16** (1971), 229–248.
7. M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320,938.
8. ______, *Probability and related topics in physical sciences*, Interscience, London, 1959.
9. S.V. Konyagin, *On minimum of modulus of random trigonometric oolynomials with coefficients* ± 1, Matem. zametki **56** (1994), 80–101, In Russian.
10. B.F. Logan and L.A. Shepp, *Real zeros of random polynomials*, Proc. London Math. Soc. Third Series **18** (1968), 29–35.
11. ______, *Real zeros of random polynomials ii*, Proc. London Math. Soc. Third Series **18** (1968), 308–314.
12. S.O. Rice, *Mathematical theory of random noise*, Bell System Tech. J. **25** (1945), 46–156.
13. D.C. Stevens, *The average and variance of the number of real zeros of random functions*, Ph.D. Dissertation, New York University, New York, 1965.
14. D.I. Šparo and M.G. Šur, *On the distribution of roots of random polynomials*, Vestn. Mosk. Univ., Ser. 1: Mat., Mekh. (1962), 40–53.

LARRY A. SHEPP, AT&T BELL LABORATORIES, MURRAY HILL, NJ 07974

E-mail address: las@research.att.com

ROBERT J. VANDERBEI, PROGRAM IN STATISTICS AND OPERATIONS RESEARCH, PRINCETON UNIVERSITY, PRINCETON, NJ 08544

E-mail address: rvdb@princeton.edu