Condensation of the Roots of Real Random Polynomials on the Real Axis

Grégory Schehr · Satya N. Majumdar

Abstract We introduce a family of real random polynomials of degree n whose coefficients a_k are symmetric independent Gaussian variables with variance $\langle a_k^2 \rangle = e^{-k^\alpha}$, indexed by a real $\alpha \geq 0$. We compute exactly the mean number of real roots $\langle N_n \rangle$ for large n. As α is varied, one finds three different phases. First, for $0 \leq \alpha < 1$, one finds that $\langle N_n \rangle \sim (\frac{2}{\pi}) \log n$. For $1 < \alpha < 2$, there is an intermediate phase where $\langle N_n \rangle$ grows algebraically with a continuously varying exponent, $\langle N_n \rangle \sim \frac{2}{\pi} \sqrt{\frac{\alpha-1}{\alpha}} n^{\alpha/2}$. And finally for $\alpha > 2$, one finds a third phase where $\langle N_n \rangle \sim n$. This family of real random polynomials thus exhibits a condensation of their roots on the real line in the sense that, for large n, a finite fraction of their roots $\langle N_n \rangle / n$ are real. This condensation occurs via a localization of the real roots around the values $\pm \exp[\frac{\alpha}{2}(k + \frac{1}{2})^{\alpha-1}]$, $1 \ll k \leq n$.

Keywords Random polynomials · Condensation phenomenon

1 Introduction

Since the early work of Bloch and Pólya [1] in the 30's, the study of random algebraic equations has now a long story [2, 3]. In the last few years, it attracted a renewed interest in the context of probability and number theory [4], as well as in the field of quantum chaos [5]. Recently, we showed that there are also interesting connections between random polynomials and persistence properties of physical systems [6, 7].

Here we consider real random polynomials, *i.e.* polynomials with real random coefficients, of degree n. While these polynomials have exactly n roots in the complex plane, the number of roots on the *real* line N_n is a random variable. One would like to characterize

G. Schehr (✉)
Laboratoire de Physique Théorique, Université de Paris-Sud, 91405 Orsay, France
e-mail: schehr@th.u-psud.fr

S.N. Majumdar
Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, Bât. 100,
91405 Orsay Cedex, France

the statistics of this random variable and a natural question is thus: what is the mean number $\langle N_n \rangle$ of real roots and how does it behave with n for large n [4]? This question has been widely studied in the past for Kac's polynomials $K_n(x) = \sum_{k=0}^{n} a_k x^k$ where a_k are independent and identically distributed (i.i.d.) random variables of finite variance $\langle a_k^2 \rangle = \sigma^2$. In that case it is well known that $\langle N_n \rangle \sim \frac{2}{\pi} \log n$, independently of σ. This result was first obtained by Kac [8] for Gaussian random variables and it was later shown to hold also for a wider class of distributions of the coefficients a_k [2, 3]. Interesting generalizations of Kac's polynomials have been studied in the literature where a_k are independent Gaussian variables but non identical, such that $\langle a_k^2 \rangle = k^{d-2}$, where $d > 0$ is a real number, leading to $\langle N_n \rangle \sim \pi^{-1}(1 + \sqrt{d/2}) \log n$ [7, 9]. Given the robustness of this asymptotic logarithmic behavior of $\langle N_n \rangle$, it is natural to search for random polynomials for which $\langle N_n \rangle$ increases faster than $\log n$, for instance algebraically.

One such instance is provided by the real Weyl polynomials $W_n(x)$ defined by

$$W_n(x) = \sum_{k=0}^{n} \epsilon_k \frac{x^k}{\sqrt{k!}}, \tag{1}$$

where ϵ_k are i.i.d. random variables of zero mean and unit variance. Thus here, $a_k = \epsilon_k/\sqrt{k!}$ and the variance is $\langle a_k^2 \rangle = 1/k!$, which for large k behaves as $\langle a_k^2 \rangle \propto e^{-k \log k}$. For these real polynomials in (1), it is known that $\langle N_n \rangle \propto n^{1/2}$. For instance, in the special case where ϵ_k are Gaussian random variables of unit variance, one has $\langle N_n \rangle \sim \frac{2}{\pi} \sqrt{n}$ [7, 10]. Another interesting and intriguing instance of real random polynomials was introduced a long time ago by Littlewood and Offord [11] who studied the random polynomials $L_n(x)$ given by

$$L_n(x) = \frac{1}{2} + \sum_{k=1}^{n} \epsilon_k \frac{x^k}{(k!)^k}, \tag{2}$$

where $\epsilon_k = \pm 1$ with equal probability. Thus in this case $a_k = \epsilon_k/(k!)^k$ and the variance is $\langle a_k^2 \rangle = 1/(k!)^{2k}$, which behaves for large k as $\langle a_k^2 \rangle \propto e^{-2k^2 \log k}$. Using algebraic methods, they showed that such polynomials $L_n(x)$ have all their roots real and therefore $\langle N_n \rangle = n$.

We thus have here two examples of real random polynomials in (1) and (2) where, at variance with Kac's polynomials, $\langle N_n \rangle$ grows algebraically with n. In the second example (2), the number of real roots is "macroscopic" in the sense that, for large n, there is a finite fraction $\langle N_n \rangle/n$ of the roots which are on the real axis. For $L_n(x)$ in (2) this fraction is exactly one. We thus say that there is a *condensation* of the roots on the real line, similar to a Bose-Einstein condensation where a finite fraction of the particles of a quantum-mechanical system (Bosons) condense into the lowest energy level. In the case of random polynomials, the roots play the role of the particles and the equivalent of the ground state is the real line.

The purpose of this paper is to understand what types of polynomials lead to this condensation phenomenon. Of course, it is very difficult to address this question for any random coefficients a_k. However, guided by the two examples above in (1) and (2), and in particular by the large k behavior of $\langle a_k^2 \rangle$, we introduce a family of random polynomials $P_n(x)$ indexed by a real $\alpha \geq 0$ defined by

$$P_n(x) = \sum_{k=0}^{n} a_k x^k, \qquad \langle a_k^2 \rangle = e^{-k^\alpha}, \tag{3}$$
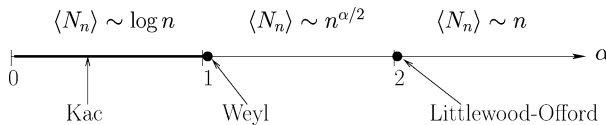
Fig. 1 Asymptotic behavior of the mean number of real roots $\langle N_n \rangle$ of $P_n(x)$ in (3) as a function of α. These polynomials exhibit a condensation of their roots on the real axis for $\alpha \geq 2$

where a_k are real independent Gaussian random variables of zero mean. While $\alpha = 0$ corresponds to Kac's polynomials, we recall that, for $W_n(x)$ in (1), $\langle a_k^2 \rangle \propto e^{-k \log k}$ and for $L_n(x)$ in (2), $\langle a_k^2 \rangle \propto e^{-2k^2 \log k}$. Therefore, due to the extra logarithmic factor, these random polynomials are not exactly of the form introduced above (3). However, for $\alpha \to 1^+$, one expects to recover the behavior of $W_n(x)$ in (1) while for $\alpha \to 2^+$, one expects $P_n(x)$ to behave similarly to $L_n(x)$ in (2): this is depicted schematically in Fig. 1.

Our main results can be summarized as follows. As $\alpha \geq 0$ is varied one finds three different *phases*. The first phase corresponds to $0 \leq \alpha < 1$, where one finds that $\langle N_n \rangle \sim (2/\pi) \log n$. In the second one, corresponding to $1 < \alpha < 2$, one has $\langle N_n \rangle \sim \frac{2}{\pi} \sqrt{\frac{\alpha-1}{\alpha}} n^{\alpha/2}$. And in the third phase, for $\alpha > 2$, one finds $\langle N_n \rangle \sim n$. The condensation of the roots on the real axis thus happens for $\alpha \geq 2$ and as one increases α, the condensation transition sets in at the critical value $\alpha_c = 2$. Furthermore, one finds that these real roots condense into a quasi-periodic structure such that there is, on average, one root in the interval $[-x_{m+1}, -x_m] \cup [x_m, x_{m+1}]$, with $x_m = e^{\frac{\alpha}{2} m^{\alpha-1}}$, with $1 \ll m < n$. These different behaviors are summarized in Fig. 1. By analogy with phase transitions of statistical systems the case $0 < \alpha < 1$ can be considered as a high-temperature phase whereas $\alpha > 2$ corresponds to the low-temperature (ordered) phase. Roughly speaking, one can consider our results as an interesting example where the transition from the high temperature where $\langle N_n \rangle \propto \log n$ (governed by a "$\alpha = 0$ fixed point") to the low temperature phase where $\langle N_n \rangle \propto n$ (governed by "$\alpha = \infty$" fixed point) happens through a *marginal phase*, for $1 < \alpha < 2$, where $\langle N_n \rangle \sim n^\varphi$ with an exponent $\varphi = \alpha/2$ which depends continuously on α.

The paper is organized as follows. In Sect. 2, we describe the general framework to compute the local density of real roots, which directly leads to $\langle N_n \rangle$. In Sects. 3 to 6 we then analyse separately the cases $0 \leq \alpha < 1$, $\alpha < 2$, $\alpha > 2$ and the "critical case" $\alpha = 2$. In Sect. 7, we give a qualitative argument to explain the condensation transition occurring at $\alpha_c = 2$ before we conclude in Sect. 8. The Appendix contains some useful technical details.

2 General Framework

First we notice that given that $P_n(x)$, as a function of x, is a Gaussian process, it is completely characterized by its two-point correlation function $C_n(x, y)$

$$C_n(x, y) = \langle P_n(x) P_n(y) \rangle = \sum_{k=0}^{n} e^{-k^\alpha} x^k y^k, \qquad (4)$$

where we used the notation $\langle \ldots \rangle$ to denote an average over the random variables a_k. A central object involved in the calculation of $\langle N_n \rangle$ is $\rho_n(x)$, the mean density of real

roots at point x. If we denote $\lambda_1, \lambda_2, \ldots, \lambda_p$ the p real roots (if any) of $P_n(x)$, one has $\delta(P_n(x)) = \sum_{i=1}^{p} \delta(x - \lambda_i)/|P_n'(\lambda_i)|$ such that $\rho_n(x)$ can be written as

$$\rho_n(x) = \sum_{i=1}^{p} \langle \delta(x - \lambda_i) \rangle = \langle |P_n'(x)| \delta(P_n(x)) \rangle$$

$$= \int_{-\infty}^{\infty} dy |y| \langle \delta(P_n'(x) - y) \delta(P_n(x)) \rangle. \tag{5}$$

Under this form (5), one observes that the computation of the mean density involves the joint distribution of the polynomial $P_n(x)$ and its derivative $P_n'(x)$ which is simply a bivariate Gaussian distribution. After Gaussian integration over y, one obtains

$$\rho_n(x) = \frac{\sqrt{c_n(x)(c_n'(x)/x + c_n''(x)) - [c_n'(x)]^2}}{2\pi c_n(x)},$$

$$c_n(x) = C_n(x, x) = \sum_{k=0}^{n} e^{-k^\alpha} x^{2k}. \tag{6}$$

This formula (6) can be written in a very compact way [4]:

$$\rho_n(x) = \frac{1}{\pi} \sqrt{\partial_u \partial_v \log C_n(u, v)}\Big|_{u=v=x}. \tag{7}$$

Given that the random coefficients a_k are drawn from a symmetric distribution, we can restrict our study of $\rho_n(x)$ on $\mathbb{R}^+$ from which one obtains the mean number of real roots $\langle N_n \rangle$ as

$$\langle N_n \rangle = 2 \int_0^{\infty} \rho_n(x) dx. \tag{8}$$

An Important Change of Variable   We will see below that it is useful to consider these polynomials $P_n(x)$ in terms of another variable Y defined as

$$Y = \left(\frac{2}{\alpha} \log x\right)^{\frac{1}{\alpha-1}}. \tag{9}$$

We denote $\hat{\rho}_n(Y)$ the mean density of the real roots in terms of this new variable such that one has also $\langle N_n \rangle = \int_0^{\infty} \hat{\rho}_n(Y) dY$. For $0 < \alpha < 1$ we will see that, for large n, most of the real roots of P_n are located close to $Y = n$ while for $\alpha > 1$, the density extends over the whole interval $Y \in [1, n]$. This change of variable (9) is motivated by the following analysis.

First we notice that $C_n(x, y) = \sum_{k=0}^{n} e^{-k^\alpha} x^k y^k$ in (4) is of the form $C_n(x, y) = c_n(\sqrt{xy})$. Anticipating a saddle point analysis, one writes $c_n(x)$ as

$$c_n(x) = \sum_{k=0}^{n} e^{-k^\alpha} x^{2k} = \sum_{k=0}^{n} \exp(-\phi(k, x)),$$

$$\phi(k, x) = k^\alpha - 2k \log x. \tag{10}$$

Although $\phi(k, x)$ is defined for integers $k = 0, 1, 2, \ldots, n$, it is readily extended to the real axis and denoted $\phi(u, x) = u^\alpha - 2u \log x$ for $u \in \mathbb{R}^+$. The behavior of $c_n(x)$ is essentially

governed by the behavior of $\phi(u, x)$ as a function of u (and fixed x). In particular, for $\alpha < 1$, $\phi(u, x)$ has a single maximum while for $\alpha > 1$, it has a single minimum for $u = u^*(x)$ given by

$$\partial_u \phi(u^*(x), x) = 0, \quad \partial_u^2 \phi(u^*(x), x) = \alpha(\alpha - 1) u^*(x)^{\alpha - 2} > 0, \tag{11}$$

$$u^*(x) = \left(\frac{2}{\alpha} \log x \right)^{\frac{1}{\alpha - 1}}. \tag{12}$$

The new variable Y introduced above in (9) is thus precisely $Y = u^*(x)$. As a consequence, the density behaves quite differently in both cases $\alpha < 1$ and $\alpha > 1$.

For $\alpha < 1$, most of the real roots on $\mathbb{R}^+$ are located in $[1, \infty]$. For fixed $x > 1$, $\phi(u, x)$ as a function of u in the interval $[0, n]$ has a global minimum for $u = n$. Therefore, the sum entering in the expression of $c_n(x)$ in (10) will be dominated by the terms with $k \sim n$. The expansion of $\phi(k, x)$ in Taylor series around $k = n$ yields

$$\phi(k, x) = \phi(n, x) + (k - n)(\alpha n^{\alpha - 1} - 2 \log x) + \cdots$$

$$= (1 - \alpha) n^\alpha - k(\alpha n^{\alpha - 1} - 2 \log x) + \cdots, \tag{13}$$

where the higher order terms can be neglected in the large n limit because $\partial^j \phi(n, x) / \partial u^j = \mathcal{O}(n^{\alpha - j})$ for $j \geq 2$. Thus, for $\alpha < 1$ one has

$$c_n(x) \sim e^{-(1-\alpha)n^\alpha} \sum_{k=0}^{n} (x e^{-\frac{\alpha}{2} n^{\alpha - 1}})^{2k}, \tag{14}$$

which, in terms of the rescaled variable $\tilde{x} = x e^{-\frac{\alpha}{2} n^{\alpha - 1}}$, is the correlator of Kac's polynomials. From this observation (14), one can straightforwardly obtain the mean number of real roots $\langle N_n \rangle$, this will be done in Sect. 3.

For $\alpha > 1$, the situation is quite different and in that case, $\phi(u, x)$ has a single minimum for $u = u^*(x) = (\frac{2}{\alpha} \log x)^{\frac{1}{\alpha - 1}}$ (11). Besides, we will see below that the main contribution to $\langle N_n \rangle$ on $\mathbb{R}^+$ comes from the interval $1 < x < \exp(\frac{\alpha}{2} n^{\alpha - 1})$ where $1 < u^*(x) < n$. In that case the sum entering in the definition of $c_n(x)$ in (10) is indeed dominated by $k \sim u^*(x)$ and $c_n(x)$ can be evaluated by a saddle point calculation. For this purpose, one obtains after some algebra explained in the Appendix, a convenient expression of $\rho_n(x)$ as

$$\rho_n(x) = \frac{1}{\pi x} \left(\frac{\sum_{k=0}^n (k - u^*(x))^2 e^{-\phi(k, x)}}{\sum_{k=0}^n e^{-\phi(k, x)}} - \left[\frac{\sum_{k=0}^n (k - u^*(x)) e^{-\phi(k, x)}}{\sum_{k=0}^n e^{-\phi(k, x)}} \right]^2 \right)^{\frac{1}{2}}, \tag{15}$$

which is the starting point of our analysis for $\alpha > 1$. For $1 < x < \exp(\frac{\alpha}{2} n^{\alpha - 1})$, one has $u^*(x) < n$ so that the sums over k in (15) are dominated by $k \sim u^*(x)$. The Taylor expansion of $\phi(k, x)$ around this minimum reads

$$\phi(k, x) = \phi(u^*(x), x) + \sum_{j=2}^{\infty} \frac{\alpha(\alpha - 1) \cdots (\alpha - j + 1)}{j!} (k - u^*(x))^j [u^*(x)]^{\alpha - j}. \tag{16}$$

For large x, $u^*(x) \propto (\log x)^{1/(\alpha - 1)}$ is also large so that, to leading order in x, one can retain only the term corresponding to $j = 2$ in the Taylor expansion in (16). This yields, for

large x

$$\sum_{k=0}^{n} g(k - u^*(x)) \exp\left(-\phi(k, x)\right)$$

$$\sim e^{-\phi(u^*(x),x)} \sum_{k=0}^{n} g(k - u^*(x)) \exp\left[-\frac{\alpha(\alpha-1)}{2}(k - u^*(x))[u^*(x)]^{\alpha-2}\right], \quad (17)$$

with $g(z) = z$ or $g(z) = z^2$ as in (15). For later purpose it is useful to write $u^*(x) = \lfloor u^*(x) \rfloor + b$ with $0 < b < 1$, where $\lfloor u^*(x) \rfloor$ is the largest integer smaller than $u^*(x)$ (*i.e.* the floor function). Performing the change of variable $m = k - \lfloor u^*(x) \rfloor$ in the discrete sum (17), such that $k - u^*(x) = m - b$ one obtains the useful expression

$$\sum_{k=0}^{n} g(k - u^*(x)) \exp\left(-\phi(k, x)\right)$$

$$\sim e^{-\phi(u^*(x),x)} \sum_{m=-\lfloor u^*(x) \rfloor}^{n-\lfloor u^*(x) \rfloor} g(m - b) \exp\left[-\frac{\alpha(\alpha-1)}{2}(m - b)^2 [u^*(x)]^{\alpha-2}\right]. \quad (18)$$

One clearly sees in expression (18) that the behavior of this discrete sum, due to the term $[u^*(x)]^{\alpha-2} \propto (\log x)^{(\alpha-2)/(\alpha-1)}$, will depend on the sign of $\alpha - 2$. We will thus treat the three cases $1 < \alpha < 2$, $\alpha > 2$ and $\alpha = 2$ separately. This will be done in Sects. 4, 5 and 6 respectively.

3 The Case $0 < \alpha < 1$

In that case, from the expression for $c_n(x)$ in (14), we can use the results of Kac's polynomials to obtain that most of the real roots will be such that, for large n, $xe^{-\frac{\alpha}{2}n^{\alpha-1}} - 1 = \mathcal{O}(n^{-1})$ [12]. In other words, the real roots are distributed in a region of width $1/n$ around $e^{\frac{\alpha}{2}n^{\alpha-1}} = 1 + \frac{\alpha}{2}n^{\alpha-1} + \mathcal{O}(n^{\alpha-2})$ and this distribution is exactly the same as the one for Kac's polynomials (corresponding to $\alpha = 0$). The number of real roots is thus also the same and given by

$$\langle N_n \rangle \sim \frac{2}{\pi} \log n, \quad (19)$$

independently of $\alpha < 1$.

4 The Case $1 < \alpha < 2$

In that case $[u^*(x)]^{\alpha-2} \to 0$ for large $u^*(x)$ and one thus sees on the asymptotic expression in (18) that the discrete sum can be replaced by an integral. This yields, for large n and large x with $x < \exp\left(\frac{\alpha}{2}n^{\alpha-1}\right)$

$$\sum_{k=0}^{n} g\left(k - u^*(x)\right) \exp\left(-\phi(k, x)\right) \sim e^{-\phi(u^*(x),x)} \int_{-\infty}^{\infty} g(y) e^{-\frac{\alpha(\alpha-1)}{2}y^2 u^*(x)^{\alpha-2}} \, dy. \quad (20)$$

Note that the prefactor $e^{-\phi(u^*(x),x)}$ is unimportant for the computation of $\rho_n(x)$ because it disappears between the numerator and the denominator in (15) and it will be omitted below.

In particular, setting $g(z) = 1$ in (20) one has

$$\sum_{k=0}^{n} \exp\left(-\phi(k, x)\right) \propto \sqrt{2\pi} \left[\frac{u^*(x)^{2-\alpha}}{\alpha(\alpha-1)}\right]^{\frac{1}{2}}, \tag{21}$$

and similarly, setting $g(z) = z^2$ in (20) one has

$$\sum_{k=0}^{n} (k - u^*(x))^2 \exp\left(-\phi(k, x)\right) \propto \sqrt{2\pi} \left[\frac{u^*(x)^{2-\alpha}}{\alpha(\alpha-1)}\right]^{\frac{3}{2}}, \tag{22}$$

while $\sum_{k=0}^{n}(k - u^*(x)) \exp\left(-\phi(k, x)\right) \sim 0$ to lowest order in n. Therefore using the exact expression given in (15) together with the asymptotic behaviors given in (21, 22), one obtains the large x behavior of $\rho_n(x)$ as

$$\rho_n(x) \sim \frac{1}{\pi x} \frac{1}{\sqrt{\alpha(\alpha-1)}} \left(\frac{2}{\alpha} \log x\right)^{\frac{2-\alpha}{2(\alpha-1)}}. \tag{23}$$

For a clear comparison with the case $\alpha > 2$, it is convenient to write the density $\hat{\rho}_n(Y)$, in terms of the variable Y, which reads, for $1 \ll Y < n$

$$\hat{\rho}_n(Y) \sim \frac{\sqrt{\alpha(\alpha-1)}}{2\pi} Y^{-\frac{1}{2}(2-\alpha)}, \tag{24}$$

and in Fig. 2(a), we show a sketch of this asymptotic behavior (24) of $\hat{\rho}_n(Y)$ for $1 \ll Y < n$.
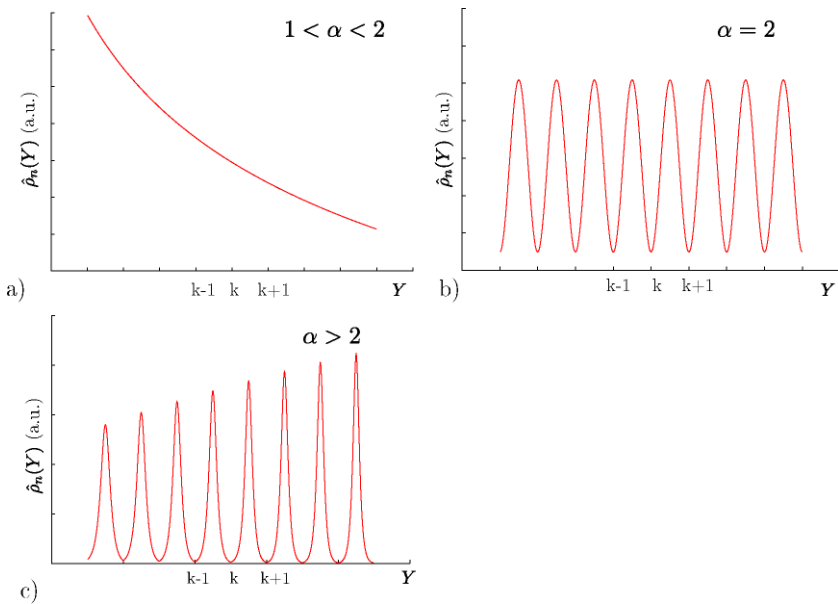

Fig. 2 (**a**) Sketch of $\hat{\rho}_n(Y)$ (in arbitrary units) given in (24) as a function of Y for $1 \ll Y < n$ for $\alpha < 2$. (**b**) Sketch of $\hat{\rho}_n(Y)$ (in arbitrary units) given in (33) as a function of Y for $1 \ll Y < n$ for $\alpha = 2$. (**c**) Sketch of $\hat{\rho}_n(Y)$ (in arbitrary units) given in (29) as a function of Y for $1 \ll Y < n$ for $\alpha > 2$. Here k denotes an integer with $1 \ll k < n$

We can now compute $\langle N_n \rangle = \int_{-\infty}^{\infty} \rho_n(x)\,dx$. First, one notices that for $\alpha > 1$, the series entering in the definition of $c_n(x)$ in (10) has an infinite radius of convergence so that one readily obtains that $\int_{-1}^{+1} \rho_n(x)\,dx$ is of order $\mathcal{O}(1)$ in the limit $n \to \infty$. Besides, for large $x \gg e^{\frac{\alpha}{2} n^{\alpha-1}}$, one has (see also [7])

$$\rho_n(x) \sim \sqrt{\frac{\langle a_{n-1}^2 \rangle}{\langle a_n^2 \rangle}} \frac{1}{\pi x^2} \sim \frac{e^{\frac{\alpha}{2} n^{\alpha-1}}}{\pi x^2}, \tag{25}$$

which implies that $\int_{e^{\frac{\alpha}{2} n^{\alpha-1}}}^{\infty} \rho_n(x)\,dx$ is also of order $\mathcal{O}(1)$ in the limit $n \to \infty$. From these properties, it follows that the main contributions to $\langle N_n \rangle$ on $\mathbb{R}^+$ comes from the interval $[1, e^{\frac{\alpha}{2} n^{\alpha-1}}]$ where the asymptotic behavior of $\rho_n(x)$ is given in (23). Therefore one has

$$\langle N_n \rangle \sim 2 \int_1^{e^{\frac{\alpha}{2} n^{\alpha-1}}} \rho_n(x)\,dx \sim \frac{2}{\pi} \sqrt{\frac{\alpha-1}{\alpha}}\, n^{\alpha/2}, \tag{26}$$

where the factor 2 comes from the additional contribution coming from $[-e^{\frac{\alpha}{2} n^{\alpha-1}}, -1]$. We thus have here an algebraic growth $\langle N_n \rangle \propto n^{\alpha/2}$ with a continuously varying exponent $\alpha/2$. This exponent tends to $1/2$ as $\alpha \to 1^+$, which is expected from the analysis of Weyl polynomials $W_n(x)$ in (1) for which $\langle a_k^2 \rangle \propto e^{-k \log k}$ (although the variance is not exactly of the form $\langle a_k^2 \rangle = e^{-k^\alpha}$). Besides, from (26), one also obtains that the amplitude of this term proportional to $n^{\alpha/2}$ vanishes when $\alpha \to 1$. We recall that for $\alpha \leq 1$, one has instead $\langle N_n \rangle \propto (\frac{2}{\pi}) \log n$ (19), characteristic for Kac's polynomials. This suggests that this limit $\alpha \to 1$ is rather singular in the sense that the asymptotic behavior of $\langle N_n \rangle$ for large n changes "discontinuously" from $\log n$ to $\sqrt{n}$.

5 The Case $\alpha > 2$

In that case, the behavior of the discrete sum in (18), which enters in the computation of $\rho_n(x)$ (15) is quite different. Indeed, in that case $[u^*(x)]^{\alpha-2} \propto (\log x)^{(\alpha-2)/(\alpha-1)} \to \infty$ for large x and therefore the leading term for large x in (18) corresponds to $m = 0$ if $b < 1/2$ or $m = 1$ in $b > 1/2$. Keeping these leading contributions, one has

$$\sum_{k=0}^{n} g\left(k - u^*(x)\right) \exp\left(-\phi(k, x)\right)$$

$$\propto g(-b) \exp\left[-\frac{\alpha(\alpha-1)}{2} b^2 u^*(x)^{\alpha-2}\right]$$

$$+ g(1-b) \exp\left[-\frac{\alpha(\alpha-1)}{2}(1-b)^2 u^*(x)^{\alpha-2}\right], \tag{27}$$

where, again, we have omitted the unimportant prefactor $e^{-\phi(u^*(x),x)}$. Using this large x expansion (27), one obtains $\rho_n(x)$ in (15) as

$$\rho_n(x) \sim \frac{2}{(\pi x) \cosh[\frac{\alpha(\alpha-1)}{2} Y^{\alpha-2}(1-2b)]}, \quad Y = \left(\frac{2}{\alpha} \log x\right)^{\frac{1}{\alpha-1}}. \tag{28}$$

In terms of the variable Y, the density $\hat{\rho}_n(Y)$ reads,

$$\hat{\rho}_n(Y = \lfloor Y \rfloor + b) \sim \frac{\alpha(\alpha - 1)Y^{\alpha-2}}{2\pi \cosh[\frac{\alpha(\alpha-1)}{2}Y^{\alpha-2}(1 - 2b)]}. \tag{29}$$

In Fig. 2(c), one shows a sketch of $\hat{\rho}_n(Y)$ for large $Y < n$ given by (29): it is qualitatively very different from the case $\alpha < 2$ (see Fig. 2(a)). Indeed, $\hat{\rho}_n(Y)$ exhibits peaks centered around $k + \frac{1}{2}$ for large integers $1 \ll k < n$. The height of these peaks is given by $\alpha(\alpha - 1)k^{\alpha-2}/(2\pi)$ whereas its width scales like $k^{2-\alpha}$.

From $\rho_n(x)$, one can now compute the mean number of real roots. As in the case $\alpha < 2$ (see (25) and above), one can show that the main contribution to $\langle N_n \rangle$ comes from the intervals $[-e^{\frac{\alpha}{2}n^{\alpha-1}}, -1]$ and $[1, e^{\frac{\alpha}{2}n^{\alpha-1}}]$. One thus has from (29)

$$\langle N_n \rangle = 2 \int_0^\infty \rho_n(x)\,dx \sim 2 \int_0^n \hat{\rho}_n(Y)\,dY$$

$$\sim \sum_{k \gg 1}^n \int_0^1 \frac{\alpha(\alpha - 1)k^{\alpha-2}}{\pi \cosh\left[\frac{\alpha(\alpha-1)}{2}k^{\alpha-2}(1 - 2b)\right]}\,db \sim \sum_{k \gg 1}^n \int_{-\frac{\alpha(\alpha-1)}{2}k^{\alpha-2}}^{\frac{\alpha(\alpha-1)}{2}k^{\alpha-2}} \frac{dz}{\pi \cosh z}, \tag{30}$$

and finally

$$\langle N_n \rangle \sim n, \tag{31}$$

where we have used $\int_{-\infty}^\infty dz/\cosh z = \pi$. This condensation of the roots on the real axis, characterized by the fact that $\langle N_n \rangle \sim n$ thus occurs via the formation of this quasi-periodic structure (see Fig. 2(c)). More precisely, this computation in (30) shows that for large k, $2 \int_k^{k+1} \hat{\rho}_n(Y)\,dY \sim 1$ which means, going back to the original variable x, that there is, on average, one root in the interval $[-x_{k+1}, -x_k] \cup [x_k, x_{k+1}]$, with $x_k = e^{\frac{\alpha}{2}k^{\alpha-1}}$.

6 The Special Case $\alpha = 2$

In view of the previous analysis, it is tempting to consider the fraction of real roots $\Phi = \lim_{n\to\infty} \langle N_n \rangle/n$ as an "order parameter". For $\alpha < 2$, one has $\Phi = 0$ whereas $\Phi = 1$ for $\alpha > 2$. One can however interpolate smoothly between these two limiting cases by considering the case $\alpha = 2$ and introducing an additional real parameter μ such that

$$\langle a_k^2 \rangle = e^{-\mu k^2}. \tag{32}$$

Performing the same algebra as explained in the Appendix, one obtains the same formula as given in (15) with $u^*(x) = \mu^{-1}\log x$. The new variable is thus here $Y = \mu^{-1}\log x$ and, setting $Y = \lfloor Y \rfloor + b$ it is easy to see that the density $\hat{\rho}_n(Y)$ is given by for $1 \ll Y < n$

$$\hat{\rho}_n(Y) = \frac{\mu}{\pi}\left[\frac{\sum_{m=-\infty}^{\infty}(m-b)^2 e^{-\mu(m-b)^2}}{\sum_{m=-\infty}^{\infty} e^{-\mu(m-b)^2}} - \left[\frac{\sum_{m=-\infty}^{\infty}(m-b)e^{-\mu(m-b)^2}}{\sum_{m=-\infty}^{\infty} e^{-\mu(m-b)^2}}\right]^2\right]^{1/2}, \tag{33}$$

which is thus 1-periodic for all μ. In Fig. 2(c), one shows a sketch of $\hat{\rho}(Y)$ for $\alpha = 2$ given by (33). For $\mu \to 0$, the density is almost constant and $\hat{\rho}_n(Y) \sim \pi^{-1}\sqrt{\mu/2}$ and the modulation of the density is increasing with μ. For large μ, the sum in (33) is dominated by

the terms corresponding to $m = 0$ and $m = 1$ and $\hat{\rho}_n(Y)$ is thus given by a formula similar to (29) setting $\alpha = 2$ and replacing $Y^{\alpha-2}$ by μ. For the average number of real roots one has

$$\langle N_n \rangle \propto \begin{cases} \frac{\sqrt{2\mu}}{\pi}n, & \mu \ll 1, \\ n, & \mu \gg 1, \end{cases}$$

which shows that this family of real random polynomials (32) interpolate smoothly between the cases $\alpha < 2$ (26) and $\alpha > 2$ (30).

7 A Qualitative Argument for the Transition at $\alpha = 2$

This condensation of the roots on the real axis can be qualitatively understood if one considers the random polynomials (for $x > 0$) $\hat{P}_n(Y) = P_n(x)$ of the variable Y, which one writes as

$$\hat{P}_n(Y) = \sum_{k=0}^{n} \hat{a}_k w(k, Y), \quad w(k, Y) = \exp\left[-\frac{1}{2}(k^\alpha - \alpha k Y^{\alpha-1})\right], \tag{34}$$

and $\hat{a}_k$ are i.i.d. Gaussian variables of unit variance. It is easy to see that the weights $w(k, Y)$, as a function of k, have a single maximum for $k = Y$ where the second derivative is proportional to $k^{\alpha-2}$. Thus for $\alpha > 2$, the weights get more and more peaked around this maximum for large k, whereas $\hat{a}_k$ is typically of order $\mathcal{O}(1)$. Therefore, given a large integer m, $\hat{P}_n(m)$ is, for $\alpha > 2$, dominated by a single term corresponding to $k = m$. Consequently, the sign of $\hat{P}_n(m)$ is essentially the sign of $\hat{a}_m$. This in turn implies that, if $\hat{a}_m$ and $\hat{a}_{m+1}$ have an opposite sign, $P_n(x)$ has, with a probability close to 1, a root in the interval $[e^{\frac{\alpha}{2}m^{\alpha-1}}, e^{\frac{\alpha}{2}(m+1)^{\alpha-1}}]$. In the case where $\hat{a}_m$ and $\hat{a}_{m+1}$ have the same sign, the same argument shows that $P_n(x)$ has, with a probability close to 1, a root in the interval $[-e^{\frac{\alpha}{2}(m+1)^{\alpha-1}}, -e^{\frac{\alpha}{2}(m)^{\alpha-1}}]$. One thus recovers qualitatively the result we had found from the computation of $\hat{\rho}_n(Y)$ in (30) where we have shown that $P_n(x)$ has, on average, one root in the interval $[-e^{\frac{\alpha}{2}(m+1)^{\alpha-1}}, -e^{\frac{\alpha}{2}(m)^{\alpha-1}}] \cup [e^{\frac{\alpha}{2}m^{\alpha-1}}, e^{\frac{\alpha}{2}(m+1)^{\alpha-1}}]$. This shows finally that $P_n(x)$ has, on average, $\langle N_n \rangle \propto n$ real roots.

We also point out that our argument explains in a rather intuitive way the result obtained by Littlewood and Offord [11] for the random polynomials $L_n(x)$ (2). For these specific polynomials, defining $x_0 = 0$, $x_m = m^m m!$, they rigorously proved, using algebraic (and rather cumbersome) methods, that $L_n(x)$ has a root either on $[x_m, x_{m+1}]$ if $\epsilon_m \epsilon_{m+1} = -1$ or in $[-x_{m+1}, -x_m]$ if $\epsilon_m \epsilon_{m+1} = 1$. Our argument gives some insight on their intriguing result and allows to understand it in a rather simple way.

8 Conclusion

To conclude we have introduced a new family of random polynomials (3), indexed by a real α. For these random polynomials, we have computed the mean density of real roots $\rho_n(x)$ from which we computed the mean number of real roots $\langle N_n \rangle$ for large n. We have shown that, while for $0 \leq \alpha < 1$, $\langle N_n \rangle \sim (\frac{2}{\pi}) \log n$, the behavior of $\langle N_n \rangle$ for $\alpha > 1$ deviates significantly from the logarithmic behavior characteristic for Kac's polynomials. For $1 < \alpha < 2$, we have shown that $\langle N_n \rangle \propto n^{\alpha/2}$ whereas for $\alpha > 2$, $\langle N_n \rangle \sim n$. This family of real random polynomials thus displays an interesting condensation phenomenon of their roots on

the real axis, which is accompanied by an ordering of the roots in a quasi periodic structure: this is depicted in Fig. 2.

Of course, the occurrence of this transition raises several interesting questions like the behavior of the variance of the number of real roots for large n as α is varied. It would be also interesting to compute the two-point correlation function of the real roots, which is a rather natural tool to characterize this periodic structure we have found. In view of this, we hope that this interesting phenomenon will stimulate further research on random polynomials.

Appendix: A Useful Expression for the Mean Density $\rho_n(x)$

In this appendix, we derive the expression for the mean density $\rho_n(x)$ as given in (15) starting from (7). We first write $c_n(x) = \langle P_n(x) P_n(x) \rangle$ as

$$c_n(x) = e^{-\phi(u^*(x),x)} \sum_{k=0}^{n} e^{-\tilde{\phi}(k,x)}, \tag{35}$$

where $u^*(x)$ is the location of the minimum of $\phi(u, x)$ given in (11)

$$u^*(x) = \left(\frac{2}{\alpha} \log x \right)^{\frac{1}{\alpha-1}}, \tag{36}$$

and

$$\phi(u^*(x), x) = (1 - \alpha) u^*(x)^{\alpha}, \tag{37}$$

$$\tilde{\phi}(k, x) = \phi(k, x) - \phi(u^*(x), x) = k^{\alpha} - \alpha k [u^*(x)]^{\alpha-1} + (\alpha - 1) [u^*(x)]^{\alpha}.$$

The correlator $C_n(x, y) = c_n(\sqrt{xy})$ is given by (35) together with (37) where x is replaced by $\sqrt{xy}$. All the dependence of $C_n(x, y)$ in x, y is thus contained in $u^*(\sqrt{xy})$ only. From its definition in (36) one has immediately

$$\partial_x u^*(\sqrt{xy}) = \frac{1}{\alpha(\alpha - 1)} \frac{1}{x} [u^*(\sqrt{xy})]^{2-\alpha}, \tag{38}$$

from which we obtain a set of useful relations

$$\partial_{x,y}^2 \phi(u^*(\sqrt{xy}), \sqrt{xy}) = -\frac{1}{\alpha(\alpha - 1)} \frac{1}{xy} [u^*(\sqrt{xy})]^{2-\alpha},$$

$$\partial_x \tilde{\phi}(k, \sqrt{xy}) = \frac{1}{x} (u^*(\sqrt{xy}) - k), \tag{39}$$

$$\partial_{x,y}^2 \tilde{\phi}(k, \sqrt{xy}) = \frac{1}{\alpha(\alpha - 1)} \frac{1}{xy} [u^*(\sqrt{xy})]^{\alpha-2}.$$

For the computation of $\rho_n(x)$ from (7), it is useful to introduce the notation, for any function $g(k)$

$$\langle g(k) \rangle_Z = \frac{\sum_{k=0}^{n} g(k) \exp(-\tilde{\phi}(k, \sqrt{xy}))}{\sum_{k=0}^{n} \exp(-\tilde{\phi}(k, \sqrt{xy}))}. \tag{40}$$

From $C_n(x, y) = c_n(\sqrt{xy})$ and $c_n(x)$ given in (35) one obtains

$$\partial_x \partial_y \log C_n(\sqrt{xy}) = -\partial_{x,y}^2 \phi(u^*(\sqrt{xy}), \sqrt{xy}) - \langle \partial_x \tilde{\phi}(k, \sqrt{xy}) \partial_y \tilde{\phi}(k, \sqrt{xy}) \rangle_Z$$
$$- \langle \partial_x \tilde{\phi}(k, \sqrt{xy}) \rangle_Z \langle \partial_x \tilde{\phi}(k, \sqrt{xy}) \rangle_Z - \langle \partial_{x,y}^2 \tilde{\phi}(k, \sqrt{xy}) \rangle_Z. \quad (41)$$

From the above relations in (39), it is readily seen that the first and the last term in (41) cancel each other. Using the relation in (7), one finally obtains the relation given in the text in (15).

References

1. Bloch, A., Pólya, G.: On the roots of certain algebraic equations. Proc. Lond. Math. Soc. (3) **33**, 102 (1932)
2. Bharucha-Reid, A.T., Sambandham, M.: Random Polynomials. Academic Press, San Diego (1986)
3. Farahmand, K.: In: Topics in Random Polynomials. Pitman Research Notes in Mathematics Series, vol. 393. Longman, Harlow (1998)
4. Edelman, A., Kostlan, E.: How many zeros of random polynomials are real? Bull. Am. Math. Soc. **32**, 1 (1995)
5. Bogomolny, E., Bohigas, O., Leboeuf, P.: Distribution of roots of random polynomials. Phys. Rev. Lett. **68**, 2726 (1992). Quantum chaotic dynamics and random polynomials. J. Stat. Phys. **85**, 639 (1996)
6. Schehr, G., Majumdar, S.N.: Statistics of the number of zero crossings: from random polynomials to the diffusion equation. Phys. Rev. Lett. **99**, 060603 (2007)
7. Schehr, G., Majumdar, S.N.: J. Stat. Phys. **132**, 235 (2008)
8. Kac, M.: On the average number of real roots of a random algebraic equation. Bull. Am. Math. Soc. **49**, 314 (1943). Erratum: Bull. Am. Math. Soc. **49**, 938 (1943)
9. Das, M.: Real zeros of a class of random algebraic polynomials. J. Indian Math. Soc. **36**, 53 (1972)
10. Leboeuf, P.: Random analytic chaotic eigenstates. J. Stat. Phys. **95**, 651 (1999)
11. Littlewood, J.E., Offord, A.C.: On the number of real roots of a random algebraic equation. II. Proc. Cambr. Philos. Soc. **35**, 133 (1939)
12. Aldous, A.P., Fyodorov, Y.V.: Real roots of random polynomials: universality close to accumulation points. J. Phys. A: Math. Gen. **37**, 1231 (2004)