数理解析研究所講究録
第 1913 巻 2014 年 93-105

Zeros of a Gaussian power series and Pfaffian †

Sho Matsumoto ‡
Graduate School of Mathematics, Nagoya University

1 Introduction

In this note, we study the random power series

$$f(z)=\sum_{k=0}^{\infty}a_{k}z^{k},$$

where the coefficients a_k are independent identically distributed (i.i.d.) standard real Gaussian random variables. In probability one, the radius of convergence of f is 1, and hence f defines a holomorphic function on the open unit disc $\mathbb{D}=\{z\in \mathbb{C}\mid|z|<1\}$. Our main purpose is to describe the distribution of zeros of f. A distribution of that kind forms a point process and can be described by its correlation functions. Our main results are to show that the correlation functions for zero distributions are given by Pfaffians, i.e., the zero distributions form Pfaffian point processes. These facts are showed independently in [5] via random matrix theory but we obtain them with a direct proof by using a Pfaffian-Hafnian identity ((2.5) below) due to Ishikawa-Kawamuko-Okada [7].

Moreover, we study the real Gaussian process $\{f(t)\}_{-1<t<1}$. We obtain the facts that the mixed moments of absolute values and those of signs

$$\mathbb{E}[|f(t_{1})f(t_{2})\cdots f(t_{n})|], \qquad \mathbb{E}[\operatorname{sgn} f(t_{1})\operatorname{sgn} f(t_{2})\cdots \operatorname{sgn} f(t_{n})]$$

can be also given by Pfaffians.

This literature is written for mathematicians other than experts on probability. In fact, we first review the fundamental knowledge for Gaussian distributions and Pfaffians in Section 2. Next, in Section 3, we see a small portion of works related to our theme: Kac's random polynomials, random power series with complex coefficients, and random matrices. Finally, our results for the random power series f are given in Section 4. We do not give the proofs of theorems in this short note. The proofs are available in the full version of the paper [12].

---

† **This is a joint work with Tomoyuki Shirai (Kyushu University).**
Workshop "Algebraic Combinatorics related to Young diagrams and Statistical Physics". August 6–10, 2012. International Institute for Advanced Study, Kyoto.

‡ Partly supported by JSPS Grant-in-Aid for Young Scientists (B) 22740060.

2 Preliminaries

2.1 Real Gaussian variables

Let $\mu_{\mathbb{R}}$ be the Lebesgue measure on the real line $\mathbb{R}$. An $\mathbb{R}$-valued random variable X is said to be *standard real Gaussian* if it has density

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad (x \in \mathbb{R}).$$

The mean (or expected value) of X is equal to 0 and the variance of X is equal to 1:

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mu_{\mathbb{R}}(dx) = 0, \qquad \mathrm{Var}[X] = \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mu_{\mathbb{R}}(dx) = 1.$$

Let $m, \sigma \in \mathbb{R}$ with $\sigma > 0$. An $\mathbb{R}$-valued random variable X is said to be a *real Gaussian with mean m and variance σ^2* if it has density

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Then we write as $X \sim \mathrm{N}_{\mathbb{R}}(m, \sigma^2)$. It is easy to confirm $\mathbb{E}[X] = m$ and $\mathrm{Var}[X] = \mathbb{E}[(X-m)^2] = \sigma^2$. If $X \sim \mathrm{N}_{\mathbb{R}}(0,1)$, then X is standard.

The *characteristic function* for $X \sim \mathrm{N}_{\mathbb{R}}(m, \sigma^2)$ is given by

$$\varphi_X(\xi) := \mathbb{E}[e^{\sqrt{-1}\xi X}] = e^{\sqrt{-1}m\xi - \frac{\sigma^2\xi^2}{2}} \qquad (\xi \in \mathbb{R}). \tag{2.1}$$

In particular, $\varphi_X(\xi) = e^{-\xi^2/2}$ if X is standard. It is well known that a distribution is uniquely determined by its characteristic function. Hence, the formula (2.1) can be regarded as a "definition" of a real Gaussian random variable.

2.2 Real Gaussian vectors

Let $\boldsymbol{m} = (m_1, \ldots, m_n)^{\mathrm{T}} \in \mathbb{R}^n$ be a column vector and let $\Sigma = (\sigma_{ij})_{1 \le i,j \le n}$ be a non-negative definite $n \times n$ real symmetric matrix. An $\mathbb{R}^n$-valued random vector $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ is said to be *real Gaussian* with parameters $(\boldsymbol{m}, \Sigma)$ if its characteristic function is given by

$$\varphi_{\boldsymbol{X}}(\boldsymbol{\xi}) := \mathbb{E}[e^{\sqrt{-1}\langle \boldsymbol{\xi}, \boldsymbol{X}\rangle}] = \exp\left(\sqrt{-1}\langle \boldsymbol{m}, \boldsymbol{\xi}\rangle - \frac{1}{2}\langle \boldsymbol{\xi}, \Sigma\boldsymbol{\xi}\rangle\right) \qquad (\boldsymbol{\xi} \in \mathbb{R}^n).$$

Here $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product on $\mathbb{R}^n$. Then we write as $\boldsymbol{X} \sim \mathrm{N}_{\mathbb{R}}(\boldsymbol{m}, \Sigma)$. It is immediate to verify

$$\mathbb{E}[X_i] = m_i \ (1 \le i \le n), \qquad \mathbb{E}[(X_i - m_i)(X_j - m_j)] = \sigma_{ij} \ (1 \le i, j \le n).$$

When Σ is positive definite, the distribution of $\boldsymbol{X}$ has the density

$$\frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}\langle \boldsymbol{x}, \Sigma^{-1}\boldsymbol{x}\rangle\right) \qquad (\boldsymbol{x} \in \mathbb{R}^n).$$

It is well known that a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ is real Gaussian if and only if linear combinations $\sum_{k=1}^{n} c_k X_k$ is real Gaussian for any $c_1, \ldots, c_n \in \mathbb{R}$.

Let Λ be a set. A family of random variables $\{X_\lambda\}_{\lambda\in\Lambda}$ is called a *real Gaussian process* if, for each $n \geq 1$ and any $\lambda_1, \ldots, \lambda_n \in \Lambda$, the random vector $(X_{\lambda_1}, \ldots, X_{\lambda_n})$ is real Gaussian.

2.3 Hafnians and Pfaffians

Let S_{2n} be the symmetric group acting on $\{1, 2, \ldots, 2n\}$ and let F_n be the subset of S_{2n} given by

$$F_n = \{\eta \in S_{2n} \mid \eta(2i-1) < \eta(2i) \ (1 \leq i \leq n), \quad \eta(1) < \eta(3) < \cdots < \eta(2n-1)\}.$$

Note that $|F_n| = (2n-1)!! = (2n-1)(2n-3)\cdots 3 \cdot 1$.

For a $2n \times 2n$ symmetric matrix $A = (a_{ij})_{1\leq i,j\leq 2n}$, the *Hafnian* of A is defined by

$$\operatorname{Hf} A = \sum_{\eta\in F_n} a_{\eta(1)\eta(2)} a_{\eta(3)\eta(4)} \cdots a_{\eta(2n-1)\eta(2n)}.$$

For a $2n \times 2n$ skew-symmetric matrix $B = (b_{ij})_{1\leq i,j\leq 2n}$, the *Pfaffian* of B is defined by

$$\operatorname{Pf} B = \sum_{\eta\in F_n} \epsilon(\eta) b_{\eta(1)\eta(2)} b_{\eta(3)\eta(4)} \cdots b_{\eta(2n-1)\eta(2n)},$$

where $\epsilon(\eta)$ is the signature of permutation η.

Example 2.1. We see the Hafnian and Pfaffian for a 4×4 symmetric matrix A and skew-symmetric matrix B. Since

$$F_2 = \left\{ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix} \right\},$$

we have

$$\operatorname{Hf} A = a_{12}a_{34} + a_{13}a_{24} + a_{14}a_{23}, \qquad \operatorname{Pf} B = b_{12}b_{34} - b_{13}b_{24} + b_{14}b_{23}.$$

2.4 Cauchy's determinants and Schur's Pfaffians

Recall important identities for determinants, Pfaffians, permanents, and Hafnian. For any $n \times n$ matrix $A = (a_{ij})_{1 \le i,j \le n}$, the *permanent* of A is defined by

$$\operatorname{per} A = \sum_{\eta \in S_n} a_{1\eta(1)} a_{2\eta(2)} \cdots a_{n\eta(n)}.$$

Lemma 2.1. *Let $x_1, x_2, \dots, y_1, y_2, \dots$ be indeterminates.*

1. *Cauchy's determinant identity:*

$$\det \left(\frac{1}{1 - x_i y_j} \right)_{1 \le i,j \le n} = \frac{\prod_{1 \le i < j \le n} (x_i - x_j) \cdot \prod_{1 \le i < j \le n} (y_i - y_j)}{\prod_{i=1}^{n} \prod_{j=1}^{n} (1 - x_i y_j)}. \tag{2.2}$$

2. *Schur's Pfaffian identity:*

$$\operatorname{Pf} \left(\frac{x_i - x_j}{1 - x_i x_j} \right)_{1 \le i,j \le 2n} = \prod_{1 \le i < j \le 2n} \frac{x_i - x_j}{1 - x_i x_j}. \tag{2.3}$$

3. *Borchardt identity:*

$$\det \left(\frac{1}{1 - x_i y_j} \right)_{1 \le i,j \le n} \cdot \operatorname{per} \left(\frac{1}{1 - x_i y_j} \right)_{1 \le i,j \le n} = \det \left(\frac{1}{(1 - x_i y_j)^2} \right)_{1 \le i,j \le n}. \tag{2.4}$$

4. *Ishikawa-Kawamuko-Okada identity:*

$$\operatorname{Pf} \left(\frac{x_i - x_j}{1 - x_i x_j} \right)_{1 \le i,j \le 2n} \cdot \operatorname{Hf} \left(\frac{1}{1 - x_i x_j} \right)_{1 \le i,j \le 2n} = \operatorname{Pf} \left(\frac{x_i - x_j}{(1 - x_i x_j)^2} \right)_{1 \le i,j \le 2n}. \tag{2.5}$$

Cauchy's determinant (2.2) is well known in combinatorics, see e.g. [11, I.4 Example 6]. Schur's Pfaffian (2.3) is a Pfaffian version of Cauchy's determinant, see [11, III.8 Example 5]. Borchardt's identity (2.4) is obtained in [1], and Carlitz and Levine [3] generalize it as follows: For an $n \times n$ matrix A with non-zero entries a_{ij} and rank ≤ 2,

$$\det \left(\frac{1}{a_{ij}} \right)_{1 \le i,j \le n} \cdot \operatorname{per} \left(\frac{1}{a_{ij}} \right)_{1 \le i,j \le n} = \det \left(\frac{1}{a_{ij}^2} \right)_{1 \le i,j \le n}.$$

Ishikawa, Kawamuko, and Okada [7] obtains (2.5), which is a Pfaffian analogue of Borchardt's identity and is the key of our proofs of theorems in Section 4.

2.5 Wick formula

Let $(X_1, \dots, X_n) \sim \mathrm{N}_{\mathbb{R}}(\mathbf{0}, \Sigma)$, where Σ is an $n \times n$ nonnegative definite real symmetric matrix. If n is odd, then we have $\mathbb{E}[X_1 X_2 \cdots X_n] = 0$. If n is even, we have

$$\mathbb{E}[X_1 X_2 \cdots X_n] = \mathrm{Hf}\, \Sigma. \tag{2.6}$$

This formula is known as the Wick formula. See an introductory survey [15].

Example 2.2. If (X_1, X_2, X_3, X_4) is mean-zero real Gaussian, then

$$\mathbb{E}[X_1 X_2 X_3 X_4] = \mathbb{E}[X_1 X_2]\mathbb{E}[X_3 X_4] + \mathbb{E}[X_1 X_3]\mathbb{E}[X_2 X_4] + \mathbb{E}[X_1 X_4]\mathbb{E}[X_2 X_3].$$

Thus, mixed moments for real Gaussian variables can be computed in a combinatorial way.

3 Random analytic functions

In this section, we shall observe a few examples of random analytic functions. General theory and recent results for random analytic functions are seen in [6].

3.1 Real zeros for random polynomials

Proposition 3.1 (Edelman-Kostlan [4]). *Let $\boldsymbol{v}(t) = (f_0(t), f_1(t), \dots, f_n(t))^{\mathrm{T}}$ be any collection of (deterministic) differentiable functions and $(a_0, a_1, \dots, a_n)$ be a real Gaussian vector with mean zero and covariance matrix Σ. Then the expected number of real solutions on a measurable set $I \subset \mathbb{R}$ of the equation*

$$a_0 f_0(t) + a_1 f_1(t) + \cdots + a_n f_n(t) = 0$$

is given by

$$\frac{1}{\pi} \int_I \left(\left[\frac{\partial^2}{\partial x \partial y} \log(\boldsymbol{v}(x)^{\mathrm{T}} \Sigma \boldsymbol{v}(y)) \right]_{x=y=t} \right)^{\frac{1}{2}} dt.$$

Example 3.1 (Kac polynomials [8]). Let $a_0, a_1, \dots, a_n$ be i.i.d. standard real Gaussian variables. A random polynomial

$$p_n(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$$

is called the Kac polynomial of degree n. We apply Proposition 3.1 with $\boldsymbol{v}(t) = (1, t, t^2, \dots, t^n)$ and $\Sigma =$ the identity matrix. Since

$$\boldsymbol{v}(x)^{\mathrm{T}} \Sigma \boldsymbol{v}(y) = 1 + xy + (xy)^2 + \cdots + (xy)^n = \frac{1 - (xy)^{n+1}}{1 - xy},$$

the expected number of real zeros of $p_n(t)$ on a measurable set $I \subset \mathbb{R}$ is

$$\frac{1}{\pi}\int_I \sqrt{\left[\frac{\partial^2}{\partial x \partial y}\log\frac{1-(xy)^{n+1}}{1-xy}\right]_{x=y=t}}\,dt = \frac{1}{\pi}\int_I \sqrt{\frac{1}{(t^2-1)^2} - \frac{(n+1)^2 t^{2n}}{(t^{2n+2}-1)^2}}\,dt.$$

In particular, this integral with $I = \mathbb{R}$ gives the expected number E_n of real zeros of p_n. It is known that

$$E_n = \frac{2}{\pi}\log n + C_1 + \frac{2}{n\pi} + O(n^{-2}) \qquad (n \to \infty)$$

with constant $C_1 = 0.6257358072\ldots$.

3.2 A random power series with random complex coefficients

If X and Y are independent standard real Gaussian variables, then we call the complex-valued random variable $Z = \frac{1}{\sqrt{2}}(X + \sqrt{-1}Y)$ a *standard complex Gaussian variable.* The density of Z with respect to the Lebesgue measure $\mu_{\mathbb{C}}$ on $\mathbb{C}$ is

$$\frac{1}{\pi}e^{-|z|^2} \qquad (z \in \mathbb{C}).$$

Note that $\mathbb{E}[Z] = 0$ and $\mathbb{E}[Z\overline{Z}] = 1$.

Consider the random power series

$$f_{\mathbb{C}}(z) = \sum_{k=0}^{\infty} \zeta_k z^k,$$

where the coefficients ζ_k are i.i.d. standard complex Gaussian. The radius of convergence is almost surely 1, and hence $f_{\mathbb{C}}$ defines a holomorphic function on the open unit disc $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ in probability one.

We consider correlation functions for zeros of $f_{\mathbb{C}}$ with respect to the Lebesgue measure $\mu_{\mathbb{C}}$. For any $n \geq 1$ and $z_1, \ldots, z_n \in \mathbb{D}$, we define the n-th *correlation function* for zeros of $f_{\mathbb{C}}$ by

$$\rho_n(z_1, \ldots, z_n) = \lim_{\epsilon \to 0} \frac{\mathrm{Prob}\{f_{\mathbb{C}} \text{ has a zero in } B_\epsilon(z_j) \text{ for each } 1 \leq j \leq n\}}{(\mu_{\mathbb{C}}(B_\epsilon))^n}.$$

Here $B_\epsilon(z)$ is the ϵ-neighborhood in $\mathbb{D}$ around z and $\mu_{\mathbb{C}}(B_\epsilon) = \mu_{\mathbb{C}}(B_\epsilon(z)) = \pi\epsilon^2$ is its volume.

Peres and Virág [14] show that the correlation functions are given in terms of determinants with the Bergman kernel:

$$\rho_n(z_1, \ldots, z_n) = \det\left(\frac{1}{(1 - z_i\overline{z_j})^2}\right)_{1 \leq i,j \leq n}. \tag{3.1}$$

Remark that they employed Borchardt's identity (2.4) in their proof.

Equation (3.1) says that the zero distribution of $f_{\mathbb{C}}$ forms a determinantal point process. As one of corollaries of the determinantal formula $\rho_n(z_1,\dots,z_n) = \det(K(z_i,z_j))_{1\le i,j\le n}$, we can obtain the following statement: The probability such that $f_{\mathbb{C}}$ has no zeros in a measurable set $C \subset \mathbb{D}$ is equal to the Fredholm determinant $\det(I - K_C)$, where K_C is the trace class operator on L^2-functions obtained by restricting K to C. Refer to the introductory survey due to Borodin [2] for determinantal point processes.

3.3 Random matrix theory

For an $N \times N$ random matrix M, its characteristic polynomial $\phi_M(\lambda) = \det(\lambda I - M)$ defines a random polynomial. Then, of course, the zeros of ϕ_M coincide with the eigenvalues of M.

We consider two kinds of random matrices. The first is a *Gaussian Orthogonal Ensemble (GOE)* matrix $M_N^{(1)}$, whose diagonal entries are $m_{ii} \sim \mathrm{N}_{\mathbb{R}}(0,1)$ and off-diagonal entries are $m_{ij} = m_{ji} \sim \mathrm{N}_{\mathbb{R}}(0,\frac{1}{2})$ $(i < j)$. The GOE matrix $M_N^{(1)}$ is an $N \times N$ real symmetric random matrix. The second is a *Gaussian Unitary Ensemble (GUE)* matrix $M_N^{(2)}$, whose diagonal entries are $m_{ii} \sim \mathrm{N}_{\mathbb{R}}(0,1)$ and off-diagonal entries are $m_{ij} = \overline{m_{ji}} \sim \mathrm{N}_{\mathbb{C}}(0,1)$ $(i < j)$. The GUE matrix $M_N^{(2)}$ is an $N \times N$ complex Hermitian random matrix. Matrix entries $\{m_{ij}\}_{1\le i\le j\le N}$ for each case are jointly independent.

The eigenvalues of both $M_N^{(1)}$ and $M_N^{(2)}$ are real. The eigenvalue density for $M_N^{(\beta)}$ $(\beta = 1,2)$ is

$$P_N^{(\beta)}(x_1,\dots,x_N) = C_N^{(\beta)} e^{-\frac{1}{2}\sum_{j=1}^N x_j^2} \prod_{1\le i<j\le N} |x_i - x_j|^\beta \qquad ((x_1,\dots,x_N) \in \mathbb{R}^N)$$

with normalization constant $C_N^{(\beta)}$, respectively.

The n-th correlation functions for eigenvalues (with respect to the Lebesgue measure $\mu_{\mathbb{R}}$) are

$$\rho_{n,N}^{(\beta)}(x_1,\dots,x_n) = \lim_{\epsilon\to 0} \frac{\mathrm{Prob}\{M_N^{(\beta)} \text{ has an eigenvalue in } (x_j-\epsilon, x_j+\epsilon) \text{ for each } 1\le j\le n\}}{(2\epsilon)^n}.$$

Here 2ϵ in the denominator means the length of the interval $(x_j - \epsilon, x_j + \epsilon)$. Using the eigenvalue density, we have

$$\rho_{n,N}^{(\beta)}(x_1,\dots,x_n) = \frac{N!}{(N-n)!} \int_{\mathbb{R}^{N-n}} P_N^{(\beta)}(x_1,\dots,x_n,x_{n+1},\dots,x_N)\,dx_{n+1}\cdots dx_N$$

for each $n = 1,2,\dots,N$, and $\rho_{n,N}^{(\beta)} \equiv 0$ for $n > N$.

It is well known that $\rho_{n,N}^{(1)}$ and $\rho_{n,N}^{(2)}$ can be given by a Pfaffian and a determinant involving Hermite polynomials, respectively. In fact, $\rho_{n,N}^{(2)}(x_1, \dots, x_n) = \det(K_N^{(2)}(x_i, x_j))_{1\le i,j\le n}$ with

$$K_N^{(2)}(x,y) = \sum_{j=0}^{N-1} \phi_j(x)\phi_j(y), \qquad \phi_j(x) = (2^j j!\sqrt{\pi})^{-1/2} e^{x^2/2}(-1)^j \frac{d^j}{dx^j} e^{-x^2}.$$

The Pfaffian expression for $\rho_{n,N}^{(1)}(x_1, \dots, x_n)$ is more complicated. See [13] for details.

4 Pfaffian expressions for correlation functions

4.1 A random power series with real coefficients

As mentioned at the beginning of the article, we consider the random power series

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

where the coefficients a_k are standard real Gaussian. The random power series f is a limiting version of Kac polynomials in Example 3.1 and a real version of random power series $f_{\mathbb{C}}(z) = \sum_{k=0}^{\infty} \zeta_k z^k$ in §3.2.

4.2 Correlations for zeros

As we did in §3.2 and §3.3, we consider correlation functions for zeros of f.

The n-th correlation function for real zeros (with respect to the Lebesgue measure $\mu_{\mathbb{R}}$) is given by

$$\rho_n^{\mathrm{r}}(t_1, \dots, t_n) = \lim_{\epsilon \to 0} \frac{\mathrm{Prob}\{f \text{ has a zero in } (t_j - \epsilon, t_j + \epsilon) \text{ for each } 1 \le j \le n\}}{(2\epsilon)^n}$$

for $t_1, t_2, \dots, t_n \in (-1, +1)$. The n-th correlation function for complex zeros (with respect to the Lebesgue measure $\mu_{\mathbb{C}}$) is given by

$$\rho_n^{\mathrm{c}}(z_1, \dots, z_n) = \lim_{\epsilon \to 0} \frac{\mathrm{Prob}\{f \text{ has a zero in } B_\epsilon(z_j) \text{ for each } 1 \le j \le n\}}{(\pi\epsilon^2)^n}$$

for $z_1, z_2, \dots, z_n \in \mathbb{D}$. We assume that imaginary parts of z_i are positive since the distribution of zeros of f is invariant under complex conjugation.

Let

$$\mathbb{K}(s,t) = \begin{pmatrix} \mathbb{K}_{11}(s,t) & \mathbb{K}_{12}(s,t) \\ \mathbb{K}_{21}(s,t) & \mathbb{K}_{22}(s,t) \end{pmatrix}$$

be the 2×2 matrix given by

$$\mathbb{K}_{11}(s,t) = \frac{s-t}{\sqrt{(1-s^2)(1-t^2)(1-st)^2}}, \quad \mathbb{K}_{12}(s,t) = \sqrt{\frac{1-t^2}{1-s^2}}\frac{1}{1-st},$$
$$\mathbb{K}_{21}(s,t) = -\sqrt{\frac{1-s^2}{1-t^2}}\frac{1}{1-st}, \quad \mathbb{K}_{22}(s,t) = \operatorname{sgn}(t-s)\arcsin\frac{\sqrt{(1-s^2)(1-t^2)}}{1-st}.$$

Here $\operatorname{sgn} t$ is

$$\operatorname{sgn} t = \begin{cases} +1, & \text{for } t > 0, \\ -1, & \text{for } t < 0, \\ 0, & \text{for } t = 0. \end{cases}$$

Our first result is to give explicit Pfaffian expressions for ρ_n^{r} and ρ_n^{c}.

Theorem 1. *The correlation function for real zeros of f is*

$$\rho_n^{\mathrm{r}}(t_1, \dots, t_n) = \pi^{-n}\operatorname{Pf}(\mathbb{K}(t_i, t_j))_{1 \le i,j \le n}.$$

Here $(\mathbb{K}(t_i, t_j))_{1 \le i,j \le n}$ stands for a $2n \times 2n$ skew-symmetric matrix by arraying 2×2 blocks $\mathbb{K}(t_i, t_j)$.

Note that entries $\mathbb{K}_{ij}(s,t)$ have the following relations

$$\mathbb{K}_{11}(s,t) = \frac{\partial^2}{\partial s \partial t}\mathbb{K}_{22}(s,t), \qquad \mathbb{K}_{12}(s,t) = -\mathbb{K}_{21}(t,s) = \frac{\partial}{\partial s}\mathbb{K}_{22}(s,t).$$

Example 4.1 (one-point correlation).

$$\rho_1^{\mathrm{r}}(t) = \pi^{-1}\mathbb{K}_{12}(t,t) = \frac{1}{\pi(1-t^2)}.$$

Therefore the expected number of real zeros of f in the interval $[a,b] \subset (-1,1)$ is

$$\int_a^b \frac{1}{\pi(1-t^2)}dt = \frac{1}{2\pi}\log\frac{(1+b)(1-a)}{(1-b)(1+a)},$$

which coincides with the limit of the result in Example 3.1.

Theorem 2. *The correlation function for complex zeros of f is*

$$\rho_n^{\mathrm{c}}(z_1, \dots, z_n) = \frac{1}{(\pi\sqrt{-1})^n}\prod_{j=1}^{n}\frac{1}{|1-z_j^2|}\cdot \operatorname{Pf}(\mathbb{K}^{\mathrm{c}}(z_i, z_j))_{1 \le i,j \le n}$$

with

$$\mathbb{K}^{\mathrm{c}}(z,w) = \begin{pmatrix} \mathrm{K}^{\mathrm{c}}(z,w) & \mathrm{K}^{\mathrm{c}}(z,\overline{w}) \\ \mathrm{K}^{\mathrm{c}}(\overline{z},w) & \mathrm{K}^{\mathrm{c}}(\overline{z},\overline{w}) \end{pmatrix}, \qquad \mathrm{K}^{\mathrm{c}}(z,w) = \frac{z-w}{(1-zw)^2}.$$

Example 4.2 (one-point correlation).

$$\rho_1^c(z) = \frac{1}{\pi\sqrt{-1}} \frac{1}{|1-z^2|} K^c(z, \bar{z}) = \frac{2\Im(z)}{\pi|1-z^2|(1-|z|^2)^2}.$$

Remark that Theorems 1 and 2 are obtained by Forrester [5] independently. In his proof, Forrester shows that the zero distribution of f coincides with eigenvalue distributions of a truncated Haar orthogonal matrix and that the eigenvalue distribution forms a Pfaffian point process. Our proof is quite different from his. In fact, in the process of our proof of Theorem 1, we obtain two new results: Theorems 3 and 4 below.

4.3 New Pfaffian identities

We next regard $\{f(t)\}_{-1<t<1}$ as a real Gaussian process. For any sequence $\boldsymbol{t} = (t_1, t_2, \dots, t_n) \in (-1, +1)^{\times n}$, the random vector $(f(t_1), f(t_2), \dots, f(t_n))$ is a real Gaussian vector with mean zero and covariance matrix $\Sigma(\boldsymbol{t}) = (\sigma(t_i, t_j))_{1 \le i,j \le n}$, where

$$\sigma(s,t) = \mathbb{E}[f(s)f(t)] = \sum_{k,l=0}^{\infty} s^k t^l \mathbb{E}[a_k a_l] = \sum_{k,l=0}^{\infty} s^k t^l \delta_{kl} = \frac{1}{1-st}.$$

For such a real Gaussian process, we obtain new Pfaffian identities.

Theorem 3. *For distinct $t_1, \dots, t_n \in (-1, +1)$, we have*

$$\mathbb{E}[|f(t_1)f(t_2)\cdots f(t_n)|] = \left(\frac{2}{\pi}\right)^{n/2} (\det \Sigma(\boldsymbol{t}))^{-1/2} \operatorname{Pf}(\mathbb{K}(t_i, t_j))_{1 \le i,j \le n}.$$

Here $\mathbb{K}(s,t)$ is defined in the previous subsection.

Note that

$$(\det \Sigma(\boldsymbol{t}))^{1/2} = \prod_{i=1}^{n} \frac{1}{\sqrt{1-t_i^2}} \cdot \prod_{1 \le i<j \le n} \frac{|t_i - t_j|}{1 - t_i t_j},$$

which follows from Cauchy's determinant identity (2.2).

Theorem 4. *For $-1 < t_1 < t_2 < \cdots < t_{2n} < 1$, we have*

$$\mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_2) \cdots \operatorname{sgn} f(t_{2n})] = \left(\frac{2}{\pi}\right)^{n} \operatorname{Pf}(\mathbb{K}_{22}(t_i, t_j))_{1 \le i,j \le 2n}. \tag{4.1}$$

Here, $\mathbb{K}_{22}(s,t) = \operatorname{sgn}(t-s) \arcsin \frac{\sqrt{(1-s^2)(1-t^2)}}{1-st}$ as before.

The formula (4.1) can be rewritten as

$$\mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_2) \cdots \operatorname{sgn} f(t_{2n})] = \operatorname{Pf}\left(\mathbb{E}[\operatorname{sgn} f(t_i) \operatorname{sgn} f(t_j)]\right)_{1 \le i < j \le 2n}$$

with $\mathbb{E}[\operatorname{sgn} f(t_i) \operatorname{sgn} f(t_j)] = \frac{2}{\pi} \arcsin \frac{\sqrt{(1-t_i^2)(1-t_j^2)}}{1-t_i t_j}$.

Example 4.3. For $-1 < t_1 < t_2 < t_3 < t_4 < 1$, we have

$$\begin{aligned}
&\mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_2) \operatorname{sgn} f(t_3) \operatorname{sgn} f(t_4)] \\
&\quad = \mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_2)] \cdot \mathbb{E}[\operatorname{sgn} f(t_3) \operatorname{sgn} f(t_4)] \\
&\qquad - \mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_3)] \cdot \mathbb{E}[\operatorname{sgn} f(t_2) \operatorname{sgn} f(t_4)] \\
&\qquad + \mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_4)] \cdot \mathbb{E}[\operatorname{sgn} f(t_2) \operatorname{sgn} f(t_3)].
\end{aligned}$$

We note that Theorem 4 implies Theorem 3 and that Theorem 3 implies Theorem 1. The complete proofs of those are seen in our full paper [12]. In our proof of Theorem 4, we show its preliminary version

$$\begin{aligned}
&\frac{\partial^{2n}}{\partial t_1 \cdots \partial t_{2n}} \mathbb{E}[\operatorname{sgn} f(t_1) \operatorname{sgn} f(t_2) \cdots \operatorname{sgn} f(t_{2n})] \\
&= \left(\frac{2}{\pi}\right)^n \prod_{i=1}^{2n} \frac{1}{\sqrt{1-t_i^2}} \cdot \operatorname{Pf}\left(\frac{t_i - t_j}{(1-t_i t_j)^2}\right)_{1 \le i,j \le 2n},
\end{aligned}$$

which is proved by using some properties for Gaussian variables and the Pfaffian-Hafnian formula (2.5).

4.4 Remarks for general covariances

Let $(X_1, \dots, X_n)$ be a real Gaussian vector with mean zero and covariance matrix $\Sigma_n = (\sigma_{ij})_{1 \le i,j \le n}$. As we saw in §2.5, mixed moments of $X_1, \dots, X_n$ is given by a Hafnian:

$$\mathbb{E}[X_1 X_2 \cdots X_n] = \begin{cases} \operatorname{Hf}(\mathbb{E}[X_i X_j])_{1 \le i,j \le n}, & \text{if } n \text{ is even,} \\ 0, & \text{if } n \text{ is odd.} \end{cases}$$

Theorem 4 is reminiscent of the Wick formula. It is natural to ask whether we can extend the formula in Theorem 4 to general cases with any covariance matrix. We first see that

$$\mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2 \cdots \operatorname{sgn} X_n] = 0 \qquad \text{if } n \text{ is odd.}$$

In fact, since $(X_1, \dots, X_n)$ and $(-X_1, \dots, -X_n)$ have the same distribution, we see that

$$\begin{aligned}
\mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2 \cdots \operatorname{sgn} X_n] &= \mathbb{E}[\operatorname{sgn}(-X_1) \operatorname{sgn}(-X_2) \cdots \operatorname{sgn}(-X_n)] \\
&= (-1)^n \mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2 \cdots \operatorname{sgn} X_n].
\end{aligned}$$

The next nontrivial case is $\mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2]$. It is not difficult to see that

$$\mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2] = \frac{2}{\pi} \arcsin \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}. \tag{4.2}$$

However, such neat formula for $\mathbb{E}[\operatorname{sgn} X_1 \operatorname{sgn} X_2 \cdots \operatorname{sgn} X_n]$ with even $n \geq 4$ is not known.

Theorem 3 gives a Pfaffian expression for moments of absolute values. Nabeya [9, 10] obtains the following formulas.

$$\mathbb{E}[|X_1X_2|] = \frac{2}{\pi}\left(\sqrt{\det \Sigma_2} + \sigma_{12} \arcsin \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}\right).$$

$$\mathbb{E}[|X_1X_2X_3|] = \left(\frac{2}{\pi}\right)^{\frac{3}{2}} \left(\det \Sigma_3 + \sum_{(i,j,k)} \frac{\sigma_{ij}\sigma_{kk} + \sigma_{ik}\sigma_{jk}}{\sqrt{\sigma_{kk}}} \arcsin \frac{\sigma_{ij}\sigma_{kk} - \sigma_{ik}\sigma_{jk}}{\sqrt{(\sigma_{ii}\sigma_{kk} - \sigma_{ik}^2)(\sigma_{jj}\sigma_{kk} - \sigma_{jk}^2)}}\right)$$

summed over $(i,j,k) = (1,2,3), (2,3,1), (3,1,2)$. However, any neat formula for moments $\mathbb{E}[|X_1X_2 \cdots X_n|]$ with $n \geq 4$ is not known.

Thus, Theorem 3 and Theorem 4 state that $\mathbb{E}[|X_1 \cdots X_n|]$ and $\mathbb{E}[\operatorname{sgn} X_1 \cdots \operatorname{sgn} X_n]$ have *surprising* Pfaffian expressions if a mean-zero real Gaussian vector $(X_1, \dots, X_n)$ has covariance matrix

$$\left(\frac{1}{1 - t_it_j}\right)_{1 \leq i,j \leq n}.$$

This result suggests the following question: *Determine covariance matrices* Σ_n *such that* $\mathbb{E}[\operatorname{sgn} X_1 \cdots \operatorname{sgn} X_n]$ *or* $\mathbb{E}[|X_1X_2 \cdots X_n|]$ *has a Pfaffian expression.*

References

[1] C. W. Borchardt, *Bestimmung der symmetrischen Verbindungen vermittelst ihrer erzeugenden Funktion*, Crelle's Journal **53** (1855), 193–198.

[2] A. Borodin, *Determinantal point processes*, arXiv:0911.1153.

[3] L. Carlitz and J. Levine, *An identity of Cayley*, American Mathematical Monthly **67** (1960), 571-573.

[4] A. Edelman and E. Kostlan, *How many zeros of a random polynomial are real?*, Bull. Amer. Math. Soc. **32** (1995), 1–37.

[5] P. J. Forrester, *The limiting Kac random polynomial and truncated random orthogonal matrices*, J. Stat. Mech. vol. 2010 P12018 (12pp) (2010).

[6] J. B. Hough, M. Krishnapur, Y. Peres and B. Virág, *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, University Lecture Series, 51. American Mathematical Society, Providence, RI, 2009.

[7] M. Ishikawa, H. Kawamuko, and S. Okada, *A Pfaffian-Hafnian analogue of Borchardt's identity*, Electron. J. Combin. **12** (2005), Note 9, 8 pp. (electronic).

[8] M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc. **49** (1943), 314–320.

[9] S. Nabeya, *Absolute moments in 2-dimensional normal distribution*, Ann. Inst. Statist. Math. Tokyo **3** (1951), 1–6.

[10] S. Nabeya, *Absolute moments in 3-dimensional normal distribution*, Ann. Inst. Statist. Math. Tokyo **4** (1952), 15–30.

[11] I. G. Macdonald, *Symmetric functions and Hall polynomials*, Second ed., Oxford University Press, 1995.

[12] S. Matsumoto and T. Shirai, *Correlation functions for zeros of a Gaussian power series and Pfaffians*, arXiv:1212.6108.

[13] M. L. Mehta, *Random Matrices*, Pure and Applied Mathematics (Amsterdam), 3rd ed., Vol. 142, Elsevier, 2004.

[14] Y. Peres and B. Virág, *Zeros of the i.i.d. Gaussian power series: a conformally invariant determinantal process*, Acta Math. **194** (2005), 1–35.

[15] A. Zvonkin, *Matrix integrals and map enumeration: an accessible introduction*, Mathl. Comput. Modelling Vol. **26** (1997), 281–304.