arXiv:1510.00536v1 [math.NT] 2 Oct 2015

CORRELATIONS BETWEEN REAL CONJUGATE ALGEBRAIC NUMBERS

FRIEDRICH GÖTZE, DZIANIS KALIADA, AND DMITRY ZAPOROZHETS

ABSTRACT. For $B \subset \mathbb{R}^k$ denote by $\Phi_k(Q;B)$ the number of ordered k-tuples in B of real conjugate algebraic numbers of degree $\le n$ and naive height $\le Q$. We show that

$$\Phi_k(Q;B) = \frac{(2Q)^{n+1}}{2\zeta(n+1)} \int_B \rho_k(\mathbf{x})\, d\mathbf{x} + O\left(Q^n\right), \quad Q \to \infty,$$

where the function ρ_k will be given explicitly. If $n = 2$, then an additional factor $\log Q$ appears in the reminder term.

1. Introdution

Baker and Schmidt [2] proved that the set of algebraic numbers of degree at most n forms a *regular system*: there exists a constant c_n depending on n only such that for any interval $I \subset \mathbb{R}^1$ and for all sufficiently large $Q \in \mathbb{N}$ there exist at least

$$c_n |I|\, Q^{n+1}/(\log Q)^{3n(n+1)}$$

algebraic numbers $\alpha_1, \dots, \alpha_l$ of degree at most n and height at most Q satisfying

$$|\alpha_i - \alpha_j| \ge (\log Q)^{3n(n+1)}/Q^{n+1}, \quad 1 \le i < j \le l.$$

Later Beresnevich [3] showed that the logarithmic factors can be omitted.

Beresnevich, Bernik, and Götze [4] obtained the following result about the distribution of distances between conjugate algebraic numbers. Let $n \ge 2$ and $0 < w \le \frac{n+1}{3}$. Then for all sufficiently large Q and any interval $I \subset [-\frac{1}{2}, \frac{1}{2}]$ there exist at least $\frac{1}{2} Q^{n+1-2w}|I|$ real algebraic numbers α of degree n and height $H(\alpha) \asymp_n Q$ having a real conjugate α^* such that $|\alpha - \alpha^*| \asymp_n Q^{-w}$.

In 1971 Brown and Mahler [5] introduced a natural generalization of the Farey sequences: the Farey sequence of degree n and order Q is the set of all real roots of integral polynomials of degree n and height at most Q. The distribution of the generalized Farey sequences has been investigated in [9] (see also [8], [10] for the case $n = 2$).

Namely, fix $n \ge 2$ and consider an arbitrary interval $I \subset \mathbb{R}$. Denote by $\Phi(Q;I)$ a number of algebraic numbers $\alpha \in I$ of degree at most n and height at most Q. Then we have that

$$\Phi(Q;I) = \frac{(2Q)^{n+1}}{2\zeta(n+1)} \int_I \rho(x)\, dx + O\left(Q^n \log^{l(n)} Q\right), \quad Q \to \infty, \tag{1}$$

where $\zeta(\cdot)$ denotes the the Riemann zeta function and $l(n)$ is defined by

$$l(n) = \begin{cases} 1, & n = 2, \\ 0, & n \ge 3. \end{cases} \tag{2}$$

2010 *Mathematics Subject Classification.* 11N45 (primary), 11C08, 60G55, 11R80 (secondary).

Key words and phrases. Conjugate algebraic numbers, correlations between algebraic numbers, distribution of algebraic numbers, integral polynomial, random polynomial.

Supported by CRC 701, Bielefeld University (Germany).

The limit density ρ is given by the formula

$$\rho(x) = 2^{-n-1} \int_{D_x} \left| \sum_{j=1}^{n} j t_j x^{j-1} \right| dt_1 \dots dt_n, \tag{3}$$

where the domain of integration D_x is defined by

$$D_x = \left\{ (t_1, \dots, t_n) \in \mathbb{R}^n \ : \ \max_{1 \le k \le n} |t_k| \le 1, \ |t_n x^n + \dots + t x| \le 1 \right\}.$$

If $x \in [-1 + 1/\sqrt{2}, 1 - 1/\sqrt{2}]$, then (3) can be simplified as follows:

$$\rho(x) = \frac{1}{12} \left(3 + \sum_{k=1}^{n-1} (k+1)^2 x^{2k} \right).$$

The function ρ coincides with the density of the real zeros of the random polynomial

$$G(x) = \xi_n x^n + \xi_{n-1} x^{n-1} + \dots + \xi_1 x + \xi_0, \tag{4}$$

where $\xi_0, \xi_1, \dots, \xi_n$ are independent random variables uniformly distributed on $[-1, 1]$ (see, e.g., [12]). It means that for any Borel subset $B \subset \mathbb{R}^1$,

$$\mathbb{E} N(G, B) = \int_B \rho(x)\, dx,$$

where $N(G, B)$ denotes the number of zeros of G lying in B. The real zeros of G can be considered as a random point process. Its distribution can be described by its k-point correlation functions $\rho_k(x_1, \dots, x_k), k = 1, 2, \dots, n$ (also known as joint intensities; see Section 2 for definition). The one-point correlation function ρ_1 coincides with the density ρ. The explicit formula for ρ_k has been obtained in [6] (see Section 2 for details).

The aim of this paper is to show that, like in the case $k = 1$, the correlation functions ρ_k are closely related with the joint distribution of real conjugate algebraic numbers.

2. Notations and main result

Let us start with some notation. Fix some positive integer $n \ge 2$ and $k \le n$. Denote

$$\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{R}^k.$$

We use the following notation for the elementary symmetric polynomials:

$$\sigma_i(\mathbf{x}) := \begin{cases} 1, & \text{if} \quad i = 0, \\ \sum_{1 \le j_1 < \dots < j_i \le k} x_{j_1} x_{j_2} \dots x_{j_i}, & \text{if} \quad 1 \le i \le k, \\ 0, & \text{otherwise}. \end{cases}$$

Denote by $\mathcal{P}(Q)$ the class of all integral polynomials of degree at most n and height at most Q. The cardinality of this class is $(2Q+1)^{n+1}$.

Recall that an integral polynomial is called *prime*, if it is irreducible over $\mathbb{Q}$, primitive (the greatest common divisor of its coefficients equals 1), and its leading coefficient is positive. Let $\mathcal{P}^*(Q)$ be the class of all prime polynomials from $\mathcal{P}(Q)$.

The *minimal polynomial* of an algebraic number α is a prime polynomial such that α is a root of this polynomial.

For a Borel subset $B \subset \mathbb{R}^k$ denote by $\Phi_k(Q; B)$ the number of ordered k-tuples $(\alpha_1, \alpha_2, \dots, \alpha_k) \in B$ of *distinct* real numbers such that for some $p \in \mathcal{P}^*(Q)$ it holds

$$p(\alpha_1) = \dots = p(\alpha_k) = 0.$$

Essentially $\Phi_k(Q; B)$ denotes the number of ordered k-tuples in B of conjugate algebraic numbers of degree at most n and height at most Q.

Given a function $g : \mathbb{R}^1 \to \mathbb{R}^1$ and a Borel subset $B \subset \mathbb{R}^k$ denote by $N_k(g, B)$ the number of ordered k-tuples $(x_1, x_2, \dots, x_k) \in B$ of *distinct* real numbers such that

$$g(x_1) = \cdots = g(x_k) = 0.$$

For any algebraic number its minimal polynomial is prime, and any prime polynomial is a minimal polynomial for some algebraic number. Therefore we have that

$$\Phi_k(Q; B) = \sum_{p \in \mathcal{P}^*(Q)} N_k(p, B). \tag{5}$$

Applying Fubini's theorem to the right hand side we obtain

$$\Phi_k(Q; B) = \sum_{m=0}^{\infty} m \cdot \#\{p \in \mathcal{P}^*(Q) \,:\, N_k(p, B) = m\}. \tag{6}$$

Since $N_k(p, B) \le n!/(n-k)!$, the sum in the right hand side is finite.

Now we are ready to state our main result.

Theorem 2.1. *Let B be a region in $\mathbb{R}^k$ with boundary consisting of a finite number of algebraic surfaces. Then*

$$\Phi_k(Q; B) = \frac{(2Q)^{n+1}}{2\zeta(n+1)} \int_B \rho_k(\mathbf{x})\, d\mathbf{x} + O\left(Q^n \log^{l(n)} Q\right), \quad Q \to \infty. \tag{7}$$

Here the function ρ_k is given by the formula

$$\rho_k(\mathbf{x}) = 2^{-n-1} \prod_{1 \le i < j \le k} |x_i - x_j| \int\limits_{D_{\mathbf{x}}} \prod_{i=1}^{k} \left| \sum_{j=0}^{n-k} t_j x_i^j \right| dt_0 \dots dt_{n-k}, \tag{8}$$

where the domain of integration $D_{\mathbf{x}}$ is defined by

$$D_{\mathbf{x}} = \left\{ (t_0, \dots, t_{n-k}) \in \mathbb{R}^{n-k+1} \,:\, \max_{0 \le i \le n} \left| \sum_{j=0}^{n-k} (-1)^{i-j} \sigma_{i-j}(\mathbf{x}) t_j \right| \le 1 \right\}.$$

The implicit big-O-constant in (7) depends on n, the number of the algebraic surfaces and their maximal degree only. The proof of Theorem 2.1 is given in Section 3.

Corollary 2.2. *The case $k = 1$ implies* (1).

Corollary 2.3. *If $k = n$, then* (8) *can be simplified as follows:*

$$\rho_n(\mathbf{x}) = \frac{2^{-n}}{(n+1)} \left(\frac{1}{\max_{0 \le i \le n} |\sigma_i(\mathbf{x})|} \right)^{n+1} \prod_{1 \le i < j \le n} |x_i - x_j|.$$

It has been shown in [6, Section 3] that the function ρ_k defined in (8) is a k-point correlation function of real zeros of the random polynomial G defined in (4). It means that for any Borel subset $B \subset \mathbb{R}^k$,

$$\mathbb{E} N_k(G, B) = \int_B \rho_k(\mathbf{x})\, d\mathbf{x}. \tag{9}$$

Let us derive several properties of ρ_k.

Proposition 2.4. *a) For any permutation s of length n,*

$$\rho_k(x_{s(1)}, x_{s(2)}, \dots, x_{s(k)}) = \rho_k(\mathbf{x}).$$

b) For all $\mathbf{x} \in \mathbb{R}^k$,

$$\rho_k(-\mathbf{x}) = \rho_k(\mathbf{x}).$$

c) For all $\mathbf{x} \in \mathbb{R}^k$ with non-zero coordinates,

$$\rho_k(x_1^{-1}, x_2^{-1}, \dots, x_k^{-1}) = \rho_k(\mathbf{x}) \prod_{i=1}^{k} x_i^2.$$

Proof. The first and the second properties are trivial. To prove the last one, note that for any integral irreducible polynomial $g(z)$ of degree n, the polynomial $z^n g(z^{-1})$ is also irreducible and has the same degree and height. Therefore for any Borel set $B \subset \mathbb{R}^k$ which does not contain points with zero coordinates we have

$$\Phi_k(Q; B^{-1}) = \Phi_k(Q; B),$$

where B^{-1} is defined as

$$B^{-1} := \left\{(x_1^{-1}, x_2^{-1}, \dots, x_k^{-1}) : (x_1, x_2, \dots, x_k) \in B\right\}.$$

Letting Q tend to infinity, we obtain from (7) that

$$\int_B \rho_k(\mathbf{x})\, d\mathbf{x} = \int_{B^{-1}} \rho_k(\mathbf{x})\, d\mathbf{x}.$$

Making the substitution $(x_1, \dots, x_k) \to (x_1^{-1}, \dots, x_k^{-1})$, we obtain

$$\int_B \rho_k(\mathbf{x})\, d\mathbf{x} = \int_B \left(\prod_{i=1}^{k} x_i^{-2}\right) \rho_k(x_1^{-1}, x_2^{-1}, \dots, x_k^{-1})\, d\mathbf{x}.$$

Since the class of sets B is large enough, the third property follows. □

3. Proof of Theorem 2.1

For a Borel set $A \subset \mathbb{R}^n$ denote by $\lambda^*(A)$ the number of points in A with coprime integer coordinates.

Consider a set $A_m \subset [-1,1]^{n+1}$ consisting of all points $(t_0, \dots, t_n) \in [-1,1]^{n+1}$ such that

$$N_k(t_n x^n + \dots + t_1 x + t_0, B) = m.$$

Then the number of primitive polynomials $p \in \mathcal{P}(Q)$ such that $N_k(p, B) = m$ is equal to $\lambda^*(QA_m)$. Hence it follows from the definition of a prime polynomial that

$$\left| \#\{p \in \mathcal{P}^*(Q) \,:\, N_k(p,B) = m\} - \frac{1}{2}\lambda^*(QA_m) \right| \le R_Q, \tag{10}$$

where R_Q denotes the number of reducible polynomials (over $\mathbb{Q}$) from $\mathcal{P}_Q$. Note that the factor $1/2$ arises because prime polynomials have positive leading coefficient. It is known (see [11]) that

$$R_Q = O\left(Q^n \log^{l(n)} Q\right), \quad Q \to \infty. \tag{11}$$

Combining (10) and (11) with (6), we obtain

$$\Phi_k(Q; B) = \frac{1}{2} \sum_{m=0}^{\infty} m \lambda^*(QA_m) + O\left(Q^n \log^{l(n)} Q\right), \quad Q \to \infty. \tag{12}$$

To estimate $\lambda^*(QA_m)$, we need the following lemma.

Lemma 3.1. *Consider a region $A \subset \mathbb{R}^d$, $d \ge 2$, with boundary consisting of a finite number of algebraic surfaces. Then*

$$\lambda^*(QA) = \frac{\mathrm{Vol}(A)}{\zeta(n)} Q^d + O\left(Q^{d-1} \log^{l(d)} Q\right), \quad Q \to \infty, \tag{13}$$

where the implicit constant in the big-O-notation depends on d, the number of the algebraic surfaces and their maximal degree only.

Proof. The results of this type are well-known, see, e.g., the classical monograph by Bachmann [1, pp. 436–444] (in particular, formulas (83a) and (83b) on pages 441–442). For the proof of Lemma 3.1, see [7]. □

Since the boundary of B consists of a finite number of algebraic surfaces, the same is true for A_m. Hence it follows from Lemma 3.1 that

$$\lambda^*(QA_m) = \frac{\mathrm{Vol}(A_m)}{\zeta(n+1)} Q^{n+1} + O\left(Q^n\right), \quad Q \to \infty,$$

which together with (12) implies

$$\Phi_k(Q;\Omega) = \frac{Q^{n+1}}{2\zeta(n+1)} \sum_{m=0}^{\infty} m \,\mathrm{Vol}(A_m) + O\left(Q^n \log^{l(n)} Q\right), \quad Q \to \infty. \tag{14}$$

To calculate $\sum_{m=0}^{\infty} m \,\mathrm{Vol}(A_m)$, note that

$$\mathrm{Vol}(A_m) = 2^{n+1}\, \mathbb{P}(N_k(G,B) = m),$$

where G is the random polynomial defined in (4). Hence

$$\sum_{m=0}^{\infty} m \,\mathrm{Vol}(A_m) = 2^{n+1}\, \mathbb{E} N_k(G,B). \tag{15}$$

Applying (9) finishes the proof.

Acknowledgments. The authors are grateful to Vasily Bernik for many useful discussions.

References

[1] P. Bachmann. *Die analytische Zahlentheorie*, volume 2. BG Teubner, Leipzig, 1894.
[2] A. Baker and W. Schmidt. Diophantine approximation and Hausdorff dimension. *Proc. London Math. Soc.*, 3(1):1–11, 1970.
[3] V. Beresnevich. On approximation of real numbers by real algebraic numbers. *Acta Arith.*, 90(2):97–112, 1999.
[4] V. Beresnevich, V. Bernik, and F. Götze. The distribution of close conjugate algebraic numbers. *Compos. Math.*, 146(5):1165–1179, 2010.
[5] H. Brown and K. Mahler. A generalization of Farey sequences: Some exploration via the computer. *J. Number Theory*, 3(3):364–370, 1971.
[6] F. Götze, D. Kaliada, and D. Zaporozhets. Correlation functions of real zeros of random polynomials. *Preprint, arXiv:1510.00025*, 2015.
[7] F. Götze, D. Kaliada, and D. Zaporozhets. Distribution of complex algebraic numbers. *Preprint, arXiv:1410.3623*, 2015.
[8] D. Kaliada. Distribution of real algebraic numbers of the second degree. *Vestsi NAN Belarusi. Ser. fiz.-mat. navuk*, (3):54–63, 2013. (In Russian).
[9] D. Kaliada. On the density function of the distribution of real algebraic numbers. *Preprint, arXiv:1405.1627*, 2014.
[10] D. Koleda. On the asymptotic distribution of algebraic numbers with growing naive height. *Chebyshevskii Sb.*, 16(1):191–204, 2015. (In Russian).
[11] B. L. van der Waerden. Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt. *Monatsh. Math. Phys.*, 43(1):133–147, 1936.
[12] D. Zaporozhets. Random polynomials and geometric probability. *Dokl. Akad. Nauk*, 71(1):53–57, 2005.

Friedrich Götze, Faculty of Mathematics, Bielefeld University, P. O. Box 10 01 31, 33501 Bielefeld, Germany
E-mail address: goetze@math.uni-bielefeld.de

Dzianis Kaliada, Institute of Mathematics, National Academy of Sciences of Belarus, 220072 Minsk, Belarus
E-mail address: koledad@rambler.ru

Dmitry Zaporozhets, St. Petersburg Department of Steklov Institute of Mathematics, Fontanka 27, 191011 St. Petersburg, Russia
E-mail address: zap1979@gmail.com