

PHILIPP HABEGGER

Separating roots of polynomials and the transfinite diameter

Abstract. Mahler proved a lower bound for the distance between distinct roots of a squarefree complex polynomial. We extend his result to packets of tuples of complex roots and slightly improve a numerical constant. One application of the former aspect is an upper bound for the transfinite diameter of certain star-shaped compact subsets of the complex plane.

Keywords. Root separation of complex polynomials, discriminant, transfinite diameter.

Mathematics Subject Classification: 12D10, 11C08, 30C85.

1 - Introduction

Let $\text{disc}(P)$ denote the discriminant of a non-zero polynomial P . For a non-negative real number t we define $\log^+ t = \log \max\{1, t\}$. Let $P \in \mathbb{C}[X] \setminus \{0\}$ with $P = a_0(X - z_1) \dots (X - z_N)$ where $z_1, \dots, z_N \in \mathbb{C}$. The Mahler measure $m(P)$ of P is $\log |a_0| + \sum_{j=1}^N \log^+ |z_j|$. It is convenient to define $M(P) = e^{m(P)}$.

Suppose P is squarefree and assume $z, w \in \mathbb{C}$ are distinct roots of P . Mahler proved the separation inequality

$$(1.1) \quad |z - w| > \sqrt{3} N^{-(N+2)/2} |\text{disc}(P)|^{1/2} M(P)^{-(N-1)}$$

in Theorem 2 [Mah64].

If P has integral coefficients, then $|\text{disc}(P)| \geq 1$ and so we may omit the discriminant in the inequality (1.1). In this setting, much effort has gone into improving the exponent of $M(P)$, see for example [Güt67, Rum79, BM04,

Received: May 25, 2021; accepted in revised form: October 27, 2021.

The author has received funding from the Swiss National Science Foundation project n° 200020_184623.

Eve04, BM10, BBG10, BD11, BD14] or Chapter 8.1 **[EG17]** for an overview of results.

In the current work we investigate such separation inequalities for complex polynomials. As both sides of (1.1) are invariant under scaling P by a non-zero complex number, the exponent of $M(P)$ cannot be improved in general. Rather our attention shifts to the factor $\sqrt{3}N^{-(N+2)/2}$. Moreover, instead of working with a pair of roots we consider arbitrary tuples of distinct roots. Our aim is to establish an obstruction to root clustering.

To formulate our main result we recall that the Barnes G -function satisfies $G(1) = G(2) = 1$ and $G(m+2) = 1!2! \cdots m!$ for all integers $m \geq 1$. Let N and m be integers with $N \geq m+1$ and $m \geq -1$. We define

$$(1.2) \quad \gamma_{N,m} = \frac{G(m+2)^2}{G(2m+3)} \prod_{j=1}^m (N^2 - j^2)^{m+1-j} > 0$$

and find

$$\gamma_{N,-1} = \gamma_{N,0} = 1, \quad \gamma_{N,1} = \frac{1}{12}(N^2 - 1), \quad \gamma_{N,2} = \frac{1}{8640}(N^2 - 1)^2(N^2 - 4), \quad \dots$$

We will see in (2.10) that $\gamma_{N,N-1} = N^{-N}$.

Theorem 1.1. *Let $P \in \mathbb{C}[X] \setminus \{0\}$ be a polynomial of degree $N \geq 2$ without multiple roots. Suppose $z_{1,0}, \dots, z_{1,m_1}, \dots, z_{n,0}, \dots, z_{n,m_n} \in \mathbb{C}$ are pairwise distinct roots of P where n, m_1, \dots, m_n are integers with $n \geq 0$ and $m_1 \geq -1, \dots, m_n \geq -1$. Then*

$$(1.3) \quad \begin{aligned} & \frac{1}{2} \log |\text{disc}(P)| \leq \\ & \sum_{l=1}^n \min \left\{ 0, \frac{1}{2} \log \gamma_{N,m_l} + \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| - m_l \sum_{j=0}^{m_l} \log^+ |z_{l,j}| \right\} \\ & + \frac{N}{2} \log N + (N-1)m(P). \end{aligned}$$

We may omit terms with $m_l = -1$ in the sum (1.3), the same holds for similar sums below.

Let us consider some special cases. Let P and N be as in the theorem.

For $n = 0$ our inequality states $|\text{disc}(P)|^{1/2} \leq N^{N/2} M(P)^{N-1}$ which is a result of Mahler, see Theorem 1 **[Mah64]**.

For $n = 1$ and $m_1 = 1$ we use $\gamma_{N,1} < N^2/12$ to see that our inequality implies

$$|z - w| > 2\sqrt{3}N^{-(N+2)/2} |\text{disc}(P)|^{1/2} M(P)^{-(N-1)}$$

for all distinct complex roots z and w of P .

For $n = 1$ and $m_1 = 2$ we use $\gamma_{N,2} < N^6/8640$ to obtain

$$|z - w||z - u||w - u| > 24\sqrt{15}N^{-(N+6)/2}|\text{disc}(P)|^{1/2}M(P)^{-(N-1)}$$

for all pairwise distinct complex roots z, w , and u of P . This improves Schönehage's Theorem 4 [Sch06], where, asymptotically in N , the constant $24\sqrt{15}$ is replaced by $2\sqrt{15}$.

For $n \geq 1$ and $m_1 = \dots = m_n = 1$ we obtain

$$\prod_{j=1}^n |z_j - z'_j| \geq (2\sqrt{3})^n N^{-(N+2n)/2} |\text{disc}(P)|^{1/2} M(P)^{-(N-1)}$$

for all pairwise distinct complex roots $z_1, z'_1, \dots, z_n, z'_n$ of P . See Mignotte's Theorem 1 [Mig95] for a more flexible estimate.

The case $n = 1$ and m_1 arbitrary is folklore [BM04] when the factor in front of $M(P)^{-(N-1)}$ is replaced by an unspecified constant that depends on N .

Our method of proof follows the approach laid out in Mahler's work [Mah64] which has found application in work of Mignotte [Mig95], Schönehage [Sch06], and others. The basic idea is to consider $\text{disc}(P)$ as a Vandermonde determinant and then do column operations to produce a common factor in many columns. In the current paper we factor out as far as possible, as suggested by Mignotte in Remark 2 [Mig95]. This factorization is done in Section 2 and leads us naturally to classical Schur polynomials. A novelty of our approach is that we replace Hadamard's Inequality for the absolute value of a determinant by the more general Fischer Inequality. In addition, we use a determinant calculation by Frame [Fra79].

It is useful to have an asymptotic upper bound for $\gamma_{N,m}$. To this end and for all $x \in (0, 1)$ we define

$$(1.4) \quad \chi(x) = -x \log x - x \log 4 + \frac{1+x^2}{2x} \log(1-x^2) + \log \frac{1+x}{1-x}.$$

Then χ is analytic on $(0, 1)$ and extends to a continuous function $\chi: [0, 1] \rightarrow \mathbb{R}$ by setting $\chi(0) = \chi(1) = 0$. As we will see in Section 3, the function χ is concave on $[0, 1]$. So its values are non-negative.

Theorem 1.2. *Let P, N, n, m_1, \dots, m_n , and the $z_{l,j}$ be as in Theorem 1.1.*

Set $p_l = (m_l + 1)/N$ for all $l \in \{1, \dots, n\}$. Then

$$\begin{aligned} \frac{1}{2} \log |\text{disc}(P)| \leq & \sum_{l=1}^n \min \left\{ 0, p_l \chi(p_l) \frac{N(N-1)}{2} + \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| - m_l \sum_{j=0}^{m_l} \log^+ |z_{l,j}| \right\} \\ & + (N-1)m(P) + O(N \log N), \end{aligned}$$

the constant implicit in $O(\cdot)$ is absolute and effective.

The modified function $\chi(x) + x \log x$ is analytic on $(-1, 1)$. Its Taylor expansion can be derived from (3.2) and yields $\chi(x) \leq -x \log x + (\frac{3}{2} - \log 4)x$. In the special case where each packet $z_{l,0}, \dots, z_{l,m_l}$ is contained in a disk of diameter ϵ we obtain the following estimate.

Corollary 1.3. *Let P, N, n, m_1, \dots, m_n , and the $z_{l,j}$ be as in Theorem 1.1. For each $l \in \{1, \dots, n\}$ let $\epsilon_l > 0$ and suppose $z_{l,0}, \dots, z_{l,m_l}$ lie in a closed disk in \mathbb{C} of radius ϵ_l . Set $p_l = (m_l + 1)/N$ for all $l \in \{1, \dots, n\}$, $p = p_1 + \dots + p_n$, and $\sigma = (p_1^2 + \dots + p_n^2)^{1/2}$. If $p > 0$, then*

$$\begin{aligned} \frac{1}{2} \log |\text{disc}(P)| \leq & \left(p \chi \left(\frac{\sigma^2}{p} \right) + \sigma^2 \log \left(\sum_{l=1}^n \frac{p_l^2 \epsilon_l}{\sigma^2} \right) \right) \frac{N(N-1)}{2} + (N-1)m(P) \\ & + O \left(N \left(\log N + \max_{1 \leq l \leq n} \log^+ \epsilon_l^{-1} \right) \right), \end{aligned}$$

the constant implicit in $O(\cdot)$ is absolute and effective.

Observe that $\sigma^2 \leq p^2 \leq p$ in the theorem. Theorem 1.2 and Corollary 1.3 are both proved in Section 3.

In a recent breakthrough, Dimitrov [**Dim**] proved the Schinzel–Zassenhaus Conjecture. His method used Dubinin’s Theorem [**Dub84**] to bound from above the transfinite diameter of a certain star shaped subset of \mathbb{C} that Dimitrov calls a Hedgehog. We explain here how to apply our, ultimately elementary, estimate to deduce an upper bound for the transfinite diameter. While our numerical constant is worse than Dubinin’s, we obtain additional information on the rate of convergence.

Let K be a non-empty compact subset of \mathbb{C} . For an integer $N \geq 2$ we define

$$d_N(K) = \sup_{z_1, \dots, z_N \in K} \left(\prod_{1 \leq i < j \leq N} |z_j - z_i| \right)^{2/(N(N-1))}.$$

It is well-known that $d_N(K)$ is non-increasing in N . The transfinite diameter of K is

$$d(K) = \lim_{N \rightarrow \infty} d_N(K).$$

We remark that the capacity of K is equal to the transfinite diameter of K , see Theorem 5.5.2 [Ran95].

For $n \in \mathbb{N} = \{1, 2, 3, \dots\}$ the Hedgehog with quills $a_1, \dots, a_n \in \mathbb{C}$ is

$$(1.5) \quad \mathcal{K}(a_1, \dots, a_n) = \bigcup_{l=1}^n [0, 1]a_l \quad \text{where} \quad [0, 1]a_l = \{\lambda a_l : \lambda \in [0, 1]\}.$$

Dubinin [Dub84] proved that $d(\mathcal{K}(a_1, \dots, a_n)) \leq 4^{-1/n} \max\{|a_1|, \dots, |a_n|\}$. The constant $1/4$ is best-possible.

The transfinite diameter satisfies $d(\lambda K) = |\lambda|d(K)$ for all $\lambda \in \mathbb{C}$. In addition, $\lambda \mathcal{K}(a_1, \dots, a_n) = \mathcal{K}(\lambda a_1, \dots, \lambda a_n)$. So to prove Dubinin's Theorem one may assume $\max\{|a_1|, \dots, |a_n|\} = 1$.

We bound from above the transfinite diameter of the union of a Hedgehog with n quills and a closed disk of radius $1 - 1/n$ centered at the origin.

Theorem 1.4. *Let $n \in \mathbb{N}$ and $a_1, \dots, a_n \in \mathbb{C}$ with $\max\{|a_1|, \dots, |a_n|\} = 1$. Set $\mathcal{K} = \mathcal{K}(a_1, \dots, a_n)$ and $\mathcal{S} = \mathcal{K} \cup \{z \in \mathbb{C} : |z| \leq 1 - 1/n\}$. Then*

$$\log d_N(\mathcal{K}) \leq \log d_N(\mathcal{S}) \leq -\frac{0.39}{n} + O\left(\frac{\log(nN)}{N}\right),$$

the constant implicit in $O(\cdot)$ is absolute and effective. In particular, $d(\mathcal{K}) \leq d(\mathcal{S}) \leq e^{-0.39/n}$.

Theorem 1.4 is proved in Section 4.

If $a_l = e^{2\pi\sqrt{-1}l/n}$ for all l , then the transfinite diameter of $\mathcal{K}(a_1, \dots, a_n) \cup \{z \in \mathbb{C} : |z| \leq 1 - 1/n\}$ equals

$$(1.6) \quad \left(\frac{(1 + (1 - \frac{1}{n})^n)^2}{4} \right)^{1/n},$$

by Table 5.1 [Ran95]. The expression inside $(\cdot)^{1/n}$ converges to $(1 + e^{-1})^2/4 = e^{-0.759\dots}$ as $n \rightarrow \infty$. We ask whether the bound $e^{-0.39/n}$ for $d(\mathcal{S})$ in Theorem 1.4 can be replaced by (1.6).

We use the big- O notation through this paper. For example, if g is a function defined on \mathbb{N} with values in $[0, \infty)$, then $O(g)$ represents a function $f: \mathbb{N} \rightarrow \mathbb{R}$ for which there exists $c > 0$ with $|f(n)| \leq cg(n)$ for all $n \in \mathbb{N}$. If not stated otherwise explicitly, the constant c will be understood as absolute.

2 - A Generalized Vandermonde Matrix

Let N and m be integers with $N \geq m + 1$ and $m \geq 0$. For independents X_0, \dots, X_m we define

$$(2.1) \quad A_N = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_0 & X_1 & \cdots & X_m \\ X_0^2 & X_1^2 & \cdots & X_m^2 \\ \vdots & \vdots & & \vdots \\ X_0^{N-1} & X_1^{N-1} & \cdots & X_m^{N-1} \end{pmatrix} \in \text{Mat}_{N, m+1}(\mathbb{Z}[X_0, \dots, X_m]).$$

If $z_0, \dots, z_m \in \mathbb{C}$ are pairwise distinct, then $\det A_N(\overline{z_0}, \dots, \overline{z_m})^t A_N(z_0, \dots, z_m) \neq 0$ by the Cauchy–Binet Formula. For $N = m + 1$ we recover a Vandermonde matrix.

The main result of this section is the following proposition. Here and below the notation $\bar{\cdot}$ denotes complex conjugation.

Proposition 2.1. *Let $N \geq 2$ and $m \geq 0$ be integers with $N \geq m + 1$ and let $z_0, \dots, z_m \in \mathbb{C}$ be pairwise distinct. Set $p = (m + 1)/N$ and*

$$X = \frac{1}{2} \log \det A_N(\overline{z_0}, \dots, \overline{z_m})^t A_N(z_0, \dots, z_m).$$

(i) *We have*

$$\begin{aligned} X \leq & \frac{1}{2} \log \gamma_{N, m} + \frac{m+1}{2} \log N + \sum_{0 \leq i < j \leq m} \log |z_j - z_i| \\ & + (N - (m + 1)) \sum_{j=0}^m \log^+ |z_j|. \end{aligned}$$

(ii) *We have*

$$X \leq \frac{m+1}{2} \log N + (N - 1) \sum_{j=0}^m \log^+ |z_j|.$$

(iii) *If $\max\{|z_0|, \dots, |z_m|\} \leq 1$, then we have*

$$(2.2) \quad X \leq (m + 1) \log N + \frac{m(m + 1)}{2} \log \max\{|z_0|, \dots, |z_m|\},$$

the right-hand side is taken as $\log N$ for $m = 0$.

Our approach is to factorize $\det A_N(X_0, \dots, X_m)^t A_N(Y_0, \dots, Y_m)$ into $\prod_{0 \leq i < j \leq m} (X_i - X_j)(Y_i - Y_j)$ times a polynomial, the result is recorded in Lemma 2.5. We recall some algebraic identities that arises in the classical theory of Schur polynomials. Our presentation is largely self-contained and we aim to provide elementary proofs or references at all steps.

Let $m \in \mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and let $I = (\alpha_0, \dots, \alpha_m)$ be an $(m+1)$ -tuple of non-negative and strictly increasing integers. We define

$$(2.3) \quad A_I = \begin{pmatrix} X_0^{\alpha_0} & \cdots & X_m^{\alpha_0} \\ \vdots & & \vdots \\ X_0^{\alpha_m} & \cdots & X_m^{\alpha_m} \end{pmatrix}.$$

For $j \in \mathbb{N}_0$ and $k \in \mathbb{Z}$ the complete homogeneous symmetric polynomial of degree k in j variables is

$$h_k = \sum_{\substack{a_0, \dots, a_j \in \mathbb{N}_0 \\ a_0 + \dots + a_j = k}} X_0^{a_0} \cdots X_j^{a_j} \in \mathbb{Z}[X_0, \dots, X_j].$$

Observe that $h_k = 0$ if $k < 0$.

The following lemma is sometimes referred to as the Jacobi–Trudi identity.

Lemma 2.2. *Let $I = (\alpha_0, \dots, \alpha_m)$ be as above, we define*

$$(2.4) \quad S_I = \det(h_{\alpha_i}(X_0), h_{\alpha_i-1}(X_0, X_1), \dots, h_{\alpha_i-m}(X_0, \dots, X_m))_{0 \leq i \leq m} \in \mathbb{Z}[X_0, \dots, X_m].$$

Then

$$(2.5) \quad \det A_I = S_I \prod_{0 \leq i < j \leq m} (X_j - X_i).$$

Proof. We claim that $\det A_I$ equals $\prod_{i=0}^{k-1} \prod_{j=i+1}^m (X_j - X_i)$ times

$$(2.6) \quad \det(h_{\alpha_i}(X_0), h_{\alpha_i-1}(X_0, X_1), \dots, h_{\alpha_i-k}(X_0, \dots, X_{k-1}, X_k), \dots, h_{\alpha_i-k}(X_0, \dots, X_{k-1}, X_m))_{0 \leq i \leq m}$$

for all $k \in \{0, \dots, m\}$. This lemma follows on taking $k = m$.

Our claim holds true for $k = 0$ as

$$A_I = (h_{\alpha_i}(X_0) \cdots h_{\alpha_i}(X_m))_{0 \leq i \leq m}.$$

The proof is by induction, so assume that the claim holds for $k \in \{0, \dots, m-1\}$.

The matrix in (2.6) has the form (c_0, \dots, c_m) where c_0, \dots, c_m are column vectors of length $m + 1$ with entries in $\mathbb{Z}[X_0, \dots, X_m]$. The determinant is alternating, so let us subtract the $(k + 1)$ -st column c_k from the $(k + 2)$ -nd column c_{k+1} , and then subtract the $(k + 1)$ -th column c_k from the $(k + 3)$ -rd column c_{k+2} , etc. until we have exhausted all columns. The induction hypothesis gives

$$(2.7) \quad \det A_I = \det(c_0, \dots, c_k, c_{k+1} - c_k, c_{k+2} - c_k, \dots, c_m - c_k) \prod_{i=0}^{k-1} \prod_{j=i+1}^m (X_j - X_i).$$

Let $j \in \{1, \dots, m - k\}$ and observe

$$c_{k+j} - c_k = (h_{\alpha_i - k}(X_0, \dots, X_{k-1}, X_{k+j}) - h_{\alpha_i - k}(X_0, \dots, X_k))_{0 \leq i \leq m}$$

with

$$\begin{aligned} & h_{\alpha_i - k}(X_0, \dots, X_{k-1}, X_{k+j}) - h_{\alpha_i - k}(X_0, \dots, X_k) \\ &= \sum_{a_0 + \dots + a_k = \alpha_i - k} X_0^{a_0} \dots X_{k-1}^{a_{k-1}} (X_{k+j}^{a_k} - X_k^{a_k}) \\ &= (X_{k+j} - X_k) \sum_{a_0 + \dots + a_k = \alpha_i - k} \sum_{a=0}^{a_k-1} X_0^{a_0} \dots X_{k-1}^{a_{k-1}} X_k^{a_k-1-a} X_{k+j}^a \\ &= (X_{k+j} - X_k) h_{\alpha_i - k-1}(X_0, \dots, X_k, X_{k+j}). \end{aligned}$$

So we can factor $X_{k+j} - X_k$ out of each respectively column. We insert into (2.7) and find that $\det A_I$ equals

$$\begin{aligned} & \det(c_0, \dots, c_k, h_{\alpha_i - (k+1)}(X_0, \dots, X_k, X_{k+1}), \dots, \\ & \quad h_{\alpha_i - (k+1)}(X_0, \dots, X_k, X_m))_{0 \leq i \leq m} \\ & \quad \times \left(\prod_{i=0}^{k-1} \prod_{j=i+1}^m (X_j - X_i) \right) \prod_{j=1}^{m-k} (X_{k+j} - X_k). \end{aligned}$$

The Vandermonde factor equals $\prod_{i=0}^k \prod_{j=i+1}^m (X_j - X_i)$. So we have verified (2.6) for $k + 1$. \square

We come to a further lemma, well-known from the theory of Schur polynomials.

Lemma 2.3. *Let I be as above and let S_I be as in (2.4). The coefficients of S_I are non-negative integers.*

Proof. See for example Lemma in [Pro89] for a sketch of a proof that involves only basic properties of the determinant. \square

We come to an elementary lemma on vanishing of the determinant.

Lemma 2.4. *Let $D \in \text{Mat}_{m+1}(\mathbb{C}[T])$ and $t \in \mathbb{C}$ such that the rank of $D(t) \in \text{Mat}_{m+1}(\mathbb{C})$ is at most r . Then $\det D \in \mathbb{C}[T]$ has a zero of order at least $m + 1 - r$ at t .*

Proof. As $\mathbb{C}[T]$ is a principal ideal domain we can put D into Smith normal form. In other words, there are matrices $U, V \in \text{GL}_{m+1}(\mathbb{C}[T])$ such that UDV is diagonal with diagonal entries $f_0, \dots, f_m \in \mathbb{C}[T]$. Note that $\det U$ and $\det V$ are non-zero constants. Therefore, the order of vanishing of $\det D$ at t equals the order of vanishing of $f_0 \cdots f_m$ at t . The lemma follows as by hypothesis, at most r among $f_0(t), \dots, f_m(t)$ are non-zero. \square

We subsume our results in the next lemma. The identity (2.9) follows from a computation of Frame. Recall that $\gamma_{N,m}$ was defined in (1.2).

Lemma 2.5. *Suppose $N \geq 2$ and $m \geq 0$ are integers with $N \geq m + 1$. Then*

$$(2.8) \quad \det A_N(X_0, \dots, X_m)^t A_N(Y_0, \dots, Y_m) = B \prod_{0 \leq i < j \leq m} (X_j - X_i)(Y_j - Y_i)$$

where $B \in \mathbb{Z}[X_0, \dots, X_m, Y_0, \dots, Y_m]$ has non-negative coefficients with

$$(2.9) \quad B(\underbrace{1, \dots, 1}_{2m+2 \text{ times}}) = \gamma_{N,m} N^{m+1}.$$

Finally, $\max_{0 \leq j \leq m} \{\deg_{X_j} B, \deg_{Y_j} B\} \leq N - (m + 1)$.

For $N = m + 1$ we are in the Vandermonde case and (2.9) implies

$$(2.10) \quad \gamma_{N,N-1} = N^{-N}.$$

Proof of Lemma 2.5. By the Cauchy–Binet Formula, the left-hand side of (2.8) equals

$$\sum_I \det A_I(X_0, \dots, X_m)^t A_I(Y_0, \dots, Y_m),$$

where here and below the sum ranges over all $(m + 1)$ -tuples $I = (\alpha_0, \dots, \alpha_m)$ of integers satisfying $0 \leq \alpha_0 < \dots < \alpha_m \leq N - 1$. Lemma 2.2 implies (2.8) with

$$B = \sum_I S_I(X_0, \dots, X_m) S_I(Y_0, \dots, Y_m)$$

and with S_I as in (2.4). Thus $B \in \mathbb{Z}[X_1, \dots, X_m, Y_1, \dots, Y_m]$. Moreover, each S_I has non-negative coefficients by Lemma 2.3 and thus the same holds for B .

The degree of $\det A_N(X_0, \dots, X_m)^t A_N(Y_0, \dots, Y_m)$ with respect to X_j is at most $N - 1$. The degree of the Vandermonde determinant $\prod_{0 \leq i < j \leq m} (X_j - X_i)$ with respect to X_j is m . So (2.8) implies $\deg_{X_j} B \leq N - (m + 1)$ and the same bound holds for $\deg_{Y_j} B$.

It remains to justify the value of B at $(1, \dots, 1)$. To this end, here and below I is as before. Observe that

$$S_I(1, \dots, 1) = \det(h_{\alpha_i}(1), h_{\alpha_i-1}(1, 1), \dots, h_{\alpha_i-m}(1, \dots, 1))_{0 \leq i \leq m}$$

and $h_{\alpha_i-j}(\underbrace{1, \dots, 1}_{j+1 \text{ times}}) = \binom{\alpha_i}{j}$. So

$$B(1, \dots, 1) = \sum_I b_I^2 \quad \text{where} \quad b_{(\alpha_0, \dots, \alpha_m)} = \det \left(\binom{\alpha_i}{j} \right)_{0 \leq i, j \leq m}.$$

Observe

$$b_{(\alpha_0, \dots, \alpha_m)} = \frac{1}{1!2! \dots m!} \det(\alpha_i(\alpha_i - 1) \dots (\alpha_i - j + 1))_{0 \leq i, j \leq m}.$$

A typical entry in the $(j + 1)$ -st column of the matrix is of the form $\alpha_i^j +$ (polynomial in α_i of degree $< j$). So the determinant on the right is a Vandermonde determinant in disguise; for a reference see Proposition 1 [Kra99]. So

$$b_{(\alpha_0, \dots, \alpha_m)} = \frac{1}{G(m + 2)} \det(\alpha_i^j)_{0 \leq i, j \leq m}$$

where G is the Barnes G -function and where we use the convention $0^0 = 1$. Thus

$$(2.11) \quad B(1, \dots, 1) = \sum_I b_I^2 = \frac{\det C^t C}{G(m + 2)^2} \quad \text{with} \quad C = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^m \\ \vdots & \vdots & & \vdots \\ 1 & N - 1 & \dots & (N - 1)^m \end{pmatrix}$$

by another application of the Cauchy–Binet Formula. We find the power-sum Hankel matrix

$$C^t C = \left(\sum_{k=0}^{N-1} k^{i+j} \right)_{0 \leq i, j \leq m}.$$

Note that the top-left entry is N .

Frame, see equation (1.3) [Fra79], computed

$$(2.12) \quad \det C^t C = \gamma_{N,m} N^{m+1} G(m+2)^2.$$

We will reproduce Frame's elegant argument below to keep this text as self-contained as possible. This equality together with (2.11) implies (2.9).

Let $i \geq 0$ be an integer and let $B_i = T^i + \dots \in \mathbb{Q}[T]$ denote the i -th Bernoulli polynomial with constant term $B_i(0)$. Recall that $\sum_{k=0}^{N-1} k^i = s_i(N)$ where $s_i = (B_{i+1}(T) - B_{i+1}(0))/(i+1) \in \mathbb{Q}[T]$, *e.g.*, $s_0 = T$. We define

$$D = \left(\frac{B_{i+j+1}(T) - B_{i+j+1}(0)}{i+j+1} \right)_{0 \leq i,j \leq m} = \begin{pmatrix} s_0 & s_1 & \cdots & s_m \\ s_1 & s_2 & \cdots & s_{m+1} \\ \vdots & \vdots & & \vdots \\ s_m & s_{m+1} & \cdots & s_{2m} \end{pmatrix}$$

with $D \in \text{Mat}_{m+1}(\mathbb{Q}[T])$ and find $C^t C = D(N)$. The determinant is $\det D = \sum_{\sigma} \text{sign}(\sigma) s_{0+\sigma(0)} s_{1+\sigma(1)} \cdots s_{m+\sigma(m)}$ where σ ranges over all permutations of $\{0, \dots, m\}$. As $\deg s_{i+\sigma(i)} = i + \sigma(i) + 1$ we find $\deg \det D \leq (m+1)^2$.

Clearly, $s_i(0) = 0$ and so $T \mid s_i$ for all $i \geq 0$. Therefore, $T^{m+1} \mid \det D$. Let $r \geq 1$ be an integer. The specialization

$$D(r) = \left(\sum_{k=0}^{r-1} k^{i+j} \right)_{0 \leq i,j \leq m}$$

is the product of an $(m+1) \times r$ matrix and its transpose. So its rank is at most r and Lemma 2.4 implies $(T-r)^{m+1-r} \mid \det D$ for all $r \in \{1, \dots, m\}$. Next we use the well-known symmetry $B_i(T) = (-1)^i B_i(1-T)$ for all $i \geq 0$ and $B_i(0) = 0$ for all odd $i \geq 3$ to see $s_i(1-T) = (-1)^{i+1} s_i(T)$ for all $i \geq 1$. For all $r \in \{2, \dots, m+1\}$ we see

$$-s_{i+j}(1-r) = (-1)^{i+j} s_{i+j}(r) = \sum_{k=0}^{r-1} (-k)^{i+j}$$

except when $i+j=0$ where $-s_0(1-r) = r-1$. Combining these cases gives

$$-D(1-r) = \left(\sum_{k=1}^{r-1} (-k)^{i+j} \right)_{0 \leq i,j \leq m}.$$

Note the sums are now of length $r-1$. So $-D(1-r)$ is a product of an $(m+1) \times (r-1)$ matrix with its transpose. Hence $D(1-r)$ has rank at

most $r - 1$. As above we conclude $(T + r - 1)^{m-r+2} \mid \det D$, this time for all $r \in \{2, \dots, m + 1\}$. This statement also holds true for $r = 1$ as we saw above.

We have proved that

$$\det D = \lambda \prod_{r=1}^m (T - r)^{m+1-r} \prod_{r=1}^{m+1} (T + r - 1)^{m+2-r}$$

with $\lambda \in \mathbb{Q}[T]$. Comparing degrees using $\deg \det D \leq (m + 1)^2$ we see that $\lambda \in \mathbb{Q}$.

We determine λ as follows. We have $s_{i+j} = \frac{T^{i+j+1}}{i+j+1} + (\text{lower order terms in } T)$ and

$$T^{-(m+1)^2} \det D = \sum_{\sigma} \text{sign}(\sigma) (T^{-(0+\sigma(0)+1)} s_{0+\sigma(0)}) \cdots (T^{-(m+\sigma(m)+1)} s_{m+\sigma(m)}).$$

Each term in this sum is $\text{sign}(\sigma)/(i + j - 1) + (\text{terms of order } < 0 \text{ in } T)$. We conclude that λ is the determinant of the $(m + 1) \times (m + 1)$ Hilbert matrix $(1/(i + j + 1))_{0 \leq i, j \leq m}$. The value $\lambda = G(m + 2)^4 / G(2m + 3)$ was computed by Hilbert [Hil94].

The computation yields

$$\det C^t C = \det D(N) = \frac{G(m + 2)^4}{G(2m + 3)} \prod_{r=1}^m (N - r)^{m+1-r} \prod_{r=1}^{m+1} (N + r - 1)^{m+2-r}.$$

Recalling (1.2) we conclude (2.12) and therefore the lemma. \square

Proof of Proposition 2.1. For (i) we observe that Lemma 2.5 and the triangle inequality imply

$$|B(\overline{z_0}, \dots, \overline{z_m}, z_0, \dots, z_m)| \leq B(1, \dots, 1) \prod_{j=0}^m \max\{1, |z_j|\}^{2(N-(m+1))}.$$

So we find

$$2X \leq \log(\gamma_{N,m} N^{m+1}) + 2 \sum_{0 \leq i < j \leq m} \log |z_j - z_i| + 2(N - (m + 1)) \sum_{j=0}^m \log^+ |z_j|.$$

We divide by 2 to obtain the bound in part (i).

For part (ii) we use Hadamard's Inequality. Indeed, choose $N - (m + 1)$ vectors in \mathbb{C}^N that are pairwise orthonormal and orthogonal to the columns of $A_N(z_0, \dots, z_m)$ with respect to the standard Hermitian inner product on \mathbb{C}^N .

Then apply Hadamard's Inequality, Theorem 13.5.3 [Mir55], to the $N \times N$ matrix obtained by augmenting these vectors to $A_N(z_0, \dots, z_m)$. We find

$$X \leq \log \prod_{j=0}^m (1 + |z_j|^2 + \dots + |z_j|^{2(N-1)})^{1/2} \leq \frac{m+1}{2} \log N + \sum_{j=0}^m (N-1) \log^+ |z_j|,$$

as desired.

For part (iii) we recall that $e^{2X} = \sum_I |\det A_I(z_0, \dots, z_m)|^2$ by the Cauchy-Binet Formula where, as usual, I runs over tuples $(\alpha_0, \dots, \alpha_m)$ of integers with $0 \leq \alpha_0 < \alpha_1 < \dots < \alpha_m \leq N-1$. As there are $\binom{N}{m+1}$ possible I we get $e^{2X} \leq \binom{N}{m+1} \max_I |\det A_I(z_0, \dots, z_m)|^2$. Moreover,

$$\begin{aligned} |\det A_I(z_0, \dots, z_m)| &\leq (m+1)! \max_{\sigma} |z_0|^{\alpha_{\sigma(0)}} \dots |z_m|^{\alpha_{\sigma(m)}} \\ &\leq (m+1)! \max_{\sigma} \max\{|z_0|, \dots, |z_m|\}^{\alpha_{\sigma(0)} + \dots + \alpha_{\sigma(m)}} \end{aligned}$$

where σ runs over all permutations of $\{0, \dots, m\}$. Observe that

$$\alpha_0 + \dots + \alpha_{\sigma(m)} = \alpha_0 + \dots + \alpha_m \geq 0 + 1 + \dots + m = \frac{m(m+1)}{2}.$$

As $|z_j| \leq 1$ for all j we find

$$\max\{|z_0|, \dots, |z_m|\}^{\alpha_{\sigma(0)} + \dots + \alpha_{\sigma(m)}} \leq \max\{|z_0|, \dots, |z_m|\}^{m(m+1)/2}.$$

Since $\binom{N}{m+1}^{1/2} (m+1)! \leq \binom{N}{m+1} (m+1)! \leq N^{m+1}$ we conclude $X \leq (m+1) \log N + \frac{m(m+1)}{2} \log \max\{|z_0|, \dots, |z_m|\}$, as desired. \square

Before proving Theorem 1.1 we state Fischer's Inequality, a generalization of Hadamard's Inequality for determinants.

Lemma 2.6 (Fischer's Inequality). *Let $n \in \mathbb{N}$, let $m_1, \dots, m_n \geq 0$ be integers, and set $N = (m_1 + 1) + \dots + (m_n + 1)$. For each $l \in \{1, \dots, n\}$ let $M_l \in \text{Mat}_{N, m_l+1}(\mathbb{C})$ and set $M = (M_1 \dots M_n) \in \text{Mat}_N(\mathbb{C})$. Then*

$$\det \overline{M}^t M \leq \prod_{l=1}^n \det \overline{M}_l^t M_l.$$

Proof. Let $\begin{pmatrix} M' & * \\ * & M'' \end{pmatrix} \in \text{Mat}_N(\mathbb{C})$ be a positive definite Hermitian matrix with $M' \in \text{Mat}_r(\mathbb{C})$ and $M'' \in \text{Mat}_{N-r}(\mathbb{C})$. Theorem 13.5.5 [Mir55] states

$$\det \begin{pmatrix} M' & * \\ * & M'' \end{pmatrix} \leq \det(M') \det(M'').$$

If the $N \times N$ matrix is merely positive semidefinite, then adding a positive multiple of the unit matrix leads to a positive definite Hermitian matrix. So the inequality holds for all positive semidefinite matrices by continuity. Moreover, a simple induction shows that the analog inequality holds for more than 2 matrices on the diagonal. As $\overline{M}^t M$ is positive semidefinite and Hermitian we conclude

$$\det \overline{M}^t M = \det \begin{pmatrix} \overline{M}_1^t \\ \vdots \\ \overline{M}_n^t \end{pmatrix} (M_1 \cdots M_n) = \det \begin{pmatrix} \overline{M}_1^t M_1 & & * \\ & \ddots & \\ * & & \overline{M}_n^t M_n \end{pmatrix} \\ \leq \prod_{l=1}^n \det \overline{M}_l^t M_l. \quad \square$$

Hadamard's Inequality is the case $m_1 = \cdots = m_n = 0$.

Proof of Theorem 1.1. Let $a_0 \in \mathbb{C} \setminus \{0\}$ be the leading term of P . Let $z_{0,0}, \dots, z_{0,m_0}$ be complex roots of P such that $z_{0,0}, \dots, z_{0,m_0}, \dots, z_{n,0}, \dots, z_{n,m_n}$ are pairwise distinct and constitute all complex roots of P . Here $m_0 \geq -1$.

The absolute value $|\text{disc}(P)|$ of the discriminant of P equals

$$|a_0|^{2(N-1)} \det A_N(\overline{z_{0,0}}, \dots, \overline{z_{0,m_0}}, \dots, \overline{z_{n,0}}, \dots, \overline{z_{n,m_n}})^t \\ A_N(z_{0,0}, \dots, z_{0,m_0}, \dots, z_{n,0}, \dots, z_{n,m_n}),$$

where $A_N(z_{0,0}, \dots, z_{n,m_n}) \in \text{Mat}_n(\mathbb{C})$ uses notation (2.1). Fischer's Inequality, Lemma 2.6, implies

$$(2.13) \quad \frac{1}{2} \log |\text{disc}(P)| \leq (N-1) \log |a_0| + \sum_{l=0}^n X_l$$

where $X_l = \frac{1}{2} \log \det A_N(\overline{z_{l,0}}, \dots, \overline{z_{l,m_l}})^t A_N(z_{l,0}, \dots, z_{l,m_l})$ for all $l \in \{0, \dots, n\}$ with $m_l \geq 0$ and $X_l = 0$ if $m_l = -1$; we use $\gamma_{N,-1} = 1$ below.

Let $l \in \{0, \dots, n\}$ with $m_l \geq 0$. On the one hand, Proposition 2.1(i) gives

$$(2.14) \quad X_l \leq \frac{1}{2} \log \gamma_{N,m_l} + \frac{m_l + 1}{2} \log N \\ + \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| + (N - (m_l + 1)) \sum_{j=0}^{m_l} \log^+ |z_{l,j}|.$$

On the other hand part (ii) of the same proposition gives

$$(2.15) \quad X_l \leq \frac{m_l + 1}{2} \log N + (N-1) \sum_{j=0}^{m_l} \log^+ |z_{l,j}|.$$

Combining (2.14) and (2.15) yields

$$(2.16) \quad X_l \leq \min \left\{ 0, \frac{1}{2} \log \gamma_{N, m_l} + \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| - m_l \sum_{j=0}^{m_l} \log^+ |z_{l,j}| \right\} \\ + \frac{m_l + 1}{2} \log N + (N - 1) \sum_{j=0}^{m_l} \log^+ |z_{l,j}|.$$

This bound remains true if $m_l = -1$.

We sum (2.16) over all $l \in \{0, \dots, n\}$ and insert into (2.13). Finally, recall $m(P) = \log |a_0| + \sum_{l=0}^n \sum_{j=0}^{m_l} \log^+ |z_{l,j}|$ and $(m_0 + 1) + \dots + (m_n + 1) = N$ to conclude the proof. \square

3 - Proofs of Theorem 1.2 and Corollary 1.3

We recall that all implicit constants in $O(\cdot)$ are absolute unless stated otherwise.

The function χ was defined in (1.4) for $x \in (0, 1)$ and it extends to a continuous function on $[0, 1]$ with $\chi(0) = \chi(1) = 0$. For all $x \in (0, 1)$ we have

$$(3.1) \quad \chi''(x) = \frac{\log(1 - x^2)}{x^3} < 0.$$

So χ is concave on $[0, 1]$ and in particular it takes non-negative values. Moreover, using the Taylor series of $x \mapsto \log(1 - x^2)$ we find

$$\chi''(x) = - \sum_{k=1}^{\infty} \frac{x^{2k-3}}{k}$$

on $(0, 1)$. Recall that $\chi(0) = 0$ and note that $\lim_{x \rightarrow 0} (\chi(x) + x \log x)/x = 3/2 - \log 4$, this follows easily from (1.4). So

$$(3.2) \quad \chi(x) = -x \log x + \left(\frac{3}{2} - \log 4 \right) x - \sum_{k=2}^{\infty} \frac{x^{2k-1}}{k(2k-2)(2k-1)}$$

on $(0, 1)$.

We now recall well-known growth properties of the Barnes G -function.

Lemma 3.1. *Let $m \geq 1$ be an integer, then*

$$\log G(m+1) = \frac{1}{2} m^2 \log m - \frac{3}{4} m^2 + O(m \log(m+1)).$$

Proof. By definition we have $\log G(m+1) = \sum_{j=1}^{m-1} \log j!$. So we may assume $m \geq 2$.

Let a and b be integers with $a < b$ and let $f: [a, b] \rightarrow \mathbb{R}$ be a non-decreasing continuous function. We will use the inequality $f(a) + \int_a^{b-1} f(x)dx \leq \sum_{j=a}^{b-1} f(j) \leq \int_a^b f(x)dx$.

Stirling's approximation states $\log j! = j \log j - j + O(\log(j+1))$. The map $x \mapsto x \log x - x$ is non-decreasing on $x \geq 1$. So

$$-1 + \int_1^{m-1} (x \log x - x)dx \leq \sum_{j=1}^{m-1} j \log j - j \leq \int_1^m (x \log x - x)dx.$$

The lemma follows as $x \mapsto \frac{x^2}{2} \log x - \frac{3}{4}x^2$ is an anti-derivative of $x \mapsto x \log x - x$. \square

Lemma 3.2. *Let $m \geq 0$ be an integer, then*

$$\log \frac{G(m+2)^2}{G(2m+3)} \leq -(m+1)^2 \log \frac{4(m+1)}{e^{3/2}} + O(m \log(m+1)).$$

Proof. Observe that the left-hand side vanishes for $m = 0$. So we may assume $m \geq 1$. Lemma 3.1 applied to $m+1$ and $2m+2$ implies

$$\log G(m+2) = \frac{1}{2}(m+1)^2 \log(m+1) - \frac{3}{4}(m+1)^2 + O(m \log(m+1))$$

and

$$\log G(2m+3) = 2(m+1)^2(\log 2 + \log(m+1)) - 3(m+1)^2 + O(m \log(m+1)),$$

respectively. The lemma follows on taking the difference $2 \log G(m+2) - \log G(2m+3)$. \square

Lemma 3.3. *Let N and $m \geq 0$ be integers with $N \geq m+2$. Set $p = (m+1)/N$, then*

$$\begin{aligned} \sum_{j=1}^m (m+1-j) \log(N^2 - j^2) \\ \leq p\chi(p)N^2 + (m+1)^2 \log \frac{4(m+1)}{e^{3/2}} + O((m+1) \log N). \end{aligned}$$

Proof. Let a and b be integers with $a \leq b$ and suppose $f: [a, b] \rightarrow \mathbb{R}$ is a non-increasing continuous function. Then $\sum_{j=a}^b f(j) \leq f(a) + \int_a^b f(x)dx$.

Let S denote the sum in question. Clearly, $f(x) = (m+1-x)\log(N^2-x^2)$ is non-negative and non-increasing on $[0, m+1]$. So $S \leq \sum_{j=0}^{m+1} f(j)$ and

$$\begin{aligned}
 S &\leq 2(m+1)\log N + \int_0^{m+1} (m+1-x)(\log(N-x) + \log(N+x))dx \\
 &= 2(m+1)\log N + 2\log(N) \int_0^{m+1} (m+1-x)dx \\
 (3.3) \quad &+ N^2 \int_0^p (p-y)\log(1-y^2)dy
 \end{aligned}$$

after a substitution $y = x/N$. Observe that $\int_0^{m+1} (m+1-x)dx = (m+1)^2/2$.

The function

$$y \mapsto -2py + \frac{y^2}{2} + \left(py + \frac{1-y^2}{2}\right)\log(1-y^2) + p\log\frac{1+y}{1-y}$$

is an anti-derivative of $(p-y)\log(1-y^2)$. This anti-derivative vanishes at $y = 0$ and its value at $y = p < 1$ equals $p\chi(p) + p^2\log p + p^2\log 4 - \frac{3}{2}p^2$ by (1.4). This allows us to compute the integral in (3.3) and conclude the proof. \square

We can now determine an upper bound for $\gamma_{N,m}$.

Lemma 3.4. *Let $N \geq 2$ and $m \geq 0$ be integers with $N \geq m+1$. Set $p = (m+1)/N$, then*

$$(3.4) \quad \frac{1}{2}\log \gamma_{N,m} \leq p\chi(p)\frac{N(N-1)}{2} + O((m+1)\log N).$$

Proof. If $N = m+1$, then the lemma follows from $\chi(1) = 0$ and (2.10). So we may assume $N \geq m+2$. Using the definition (1.2) we write

$$\log \gamma_{N,m} = \log \frac{G(m+2)^2}{G(2m+3)} + \sum_{j=1}^m (m+1-j)\log(N^2-j^2).$$

Adding the bounds from Lemmas 3.2 and 3.3 leads to cancellation

$$\log \gamma_{N,m} \leq p\chi(p)N^2 + O((m+1)\log N).$$

Observe that $p\chi(p)N^2 = p\chi(p)(N(N-1) + N) = p\chi(p)N(N-1) + (m+1)\chi(p)$. The lemma follows as $(m+1)\chi(p)$ ends up in the error term of (3.4); indeed, the continuous function $\chi: [0, 1] \rightarrow \mathbb{R}$ is bounded from above. \square

Proof of Theorem 1.2. We may safely omit the terms with $m_l = -1$ as then $p_l = 0$ and $\chi(0) = 0$. Lemma 3.4 furnishes the estimate $\frac{1}{2} \log \gamma_{N, m_l} \leq p_l \chi(p_l) \frac{N(N-1)}{2} + O((m_l + 1) \log N)$ if $m_l \geq 0$. Theorem 1.2 follows from Theorem 1.1(i) since $(m_1 + 1) + \cdots + (m_n + 1) \leq N$. \square

Proof of Corollary 1.3. Let z_0, \dots, z_m be pairwise distinct members of the closed unit disk in \mathbb{C} . Then $\prod_{0 \leq i < j \leq m} (z_j - z_i)$ is the Vandermonde determinant $\det(z_j^i)_{0 \leq i, j \leq m}$. Hadamard's Inequality implies $\sum_{0 \leq i < j \leq m} \log |z_j - z_i| \leq \frac{m+1}{2} \log(m+1)$. After translating and rescaling we can generalize this as follows. Let z_0, \dots, z_m be pairwise distinct members of a closed disk in \mathbb{C} of radius $\epsilon > 0$. Then $\sum_{0 \leq i < j \leq m} \log |z_j - z_i| \leq \frac{m+1}{2} \log(m+1) + \frac{m(m+1)}{2} \log \epsilon$.

We apply this bound to each of the packets $z_{l,0}, \dots, z_{l,m_l}$ and use Theorem 1.2. This theorem implies that $\frac{1}{2} \log |\text{disc}(P)|$ is at most

$$\sum_{l=1}^n \left(p_l \chi(p_l) + \frac{m_l(m_l + 1)}{N(N-1)} \log \epsilon_l \right) \frac{N(N-1)}{2} + (N-1)m(P) + O(N \log N);$$

note that

$$\sum_{l=1}^n \frac{m_l + 1}{2} \log(m_l + 1) \leq \frac{N}{2} \log N.$$

Next

$$\frac{m_l(m_l + 1)}{N(N-1)} = p_l^2 - p_l \frac{(N - m_l - 1)}{N(N-1)}$$

so we can estimate

$$\frac{m_l(m_l + 1)}{N(N-1)} \log \epsilon_l \leq p_l^2 \log \epsilon_l + p_l \frac{\log^+ \epsilon_l^{-1}}{N-1}.$$

Thus $\frac{1}{2} \log |\text{disc}(P)|$ is at most

$$(3.5) \quad \sum_{l=1}^n (p_l \chi(p_l) + p_l^2 \log \epsilon_l) \frac{N(N-1)}{2} + (N-1)m(P) + O \left(N \log N + N \sum_{l=1}^n p_l \log^+ \epsilon_l^{-1} \right).$$

Recall that $0 < p = p_1 + \cdots + p_n \leq 1$. The function $x \mapsto \chi(x)$ is concave on $[0, 1]$, as we have seen just below (3.1). Jensen's Inequality yields the bound

$$\sum_{l=1}^n p_l \chi(p_l) \leq p \chi(\sigma^2/p).$$

As $x \mapsto \log x$ is concave on $(0, \infty)$, Jensen's Inequality also implies

$$\sum_{l=1}^n p_l^2 \log \epsilon_l \leq \sigma^2 \log \left(\sum_{l=1}^n p_l^2 \epsilon_l / \sigma^2 \right).$$

These two bounds inserted in (3.5) yield the desired bound for $\frac{1}{2} \log |\text{disc}(P)|$. \square

4 - Application to Hedgehogs and Stars

In this section we use the results from Section 2 to bound from above the transfinite diameter of a Hedgehog.

Lemma 4.1. *Let $m \geq 0$ and suppose z_0, \dots, z_m lie on a line segment of length ϵ . Then*

$$(4.1) \quad \prod_{0 \leq i < j \leq m} |z_j - z_i| \leq 2^m (m+1)^{(m+1)/2} \left(\frac{\epsilon}{4} \right)^{m(m+1)/2}.$$

Before providing the proof, we remark that Stieltjes computed the maximal value of the left-hand side of (4.1), see [Sch18].

Proof. The left-hand side of (4.1) is invariant under translating all z_i . If we translate appropriately and stretch by factor $4/\epsilon$, then the product is multiplied by $(4/\epsilon)^{m(m+1)/2}$. Without loss of generality assume $\epsilon = 4$, that the line segment in question equals $[-2, 2]$, and that $m \geq 1$.

For each i we have $z_i \in [-2, 2]$, hence there is $w_i \in \mathbb{C} \setminus \{0\}$ with $|w_i| = 1$ and $w_i + w_i^{-1} = z_i$. Let

$$\begin{aligned} V &= \prod_{0 \leq i < j \leq m} |z_j - z_i| = \prod_{0 \leq i < j \leq m} |w_j - w_i + w_j^{-1} - w_i^{-1}| \\ &= \prod_{0 \leq i < j \leq m} \left| \frac{(w_i - w_j)(w_i w_j - 1)}{w_i w_j} \right| = \prod_{0 \leq i < j \leq m} |(w_i - w_j)(w_i w_j - 1)|. \end{aligned}$$

Now we apply an elementary but powerful determinant formula, see Krattenthaler's Lemma 2 [Kra99]. In our case and using $|w_j| = 1$ it implies

$$V = \frac{1}{2} \left| \det \left(w_j^i + w_j^{-i} \right)_{0 \leq i, j \leq m} \right|.$$

We have $|w_j^i + w_j^{-i}| \leq 2$ for all i and j . Hadamard's Inequality implies $V \leq \frac{1}{2} 2^{m+1} (m+1)^{(m+1)/2}$. \square

Proof of Theorem 1.4. The first inequality in the theorem follows as $d_N(K) \leq d_N(L)$ for all non-empty compact subsets $K \subseteq L \subseteq \mathbb{C}$ and all $N \geq 2$. We now prove the bound for $d_N(\mathcal{S})$ where \mathcal{S} is the star $\mathcal{K}(a_1, \dots, a_n) \cup \{z \in \mathbb{C} : |z| \leq 1 - 1/n\}$.

Let $\epsilon = 1/n \in (0, 1]$. We will continue to use the symbol ϵ to emphasize its role in the proof. Our choice of ϵ is in part made by convenience. The choice $1.06/n$ leads to a slightly better numerical estimate.

Let $N \geq 2$ and suppose $z_1, \dots, z_N \in \mathcal{S}$ are pairwise distinct. Our goal is to bound

$$v = \frac{2}{N(N-1)} \log \prod_{1 \leq i < j \leq N} |z_j - z_i| = \frac{2}{N(N-1)} \log |\det(z_j^{i-1})_{1 \leq i, j \leq N}|$$

from above where the second equality follows as the matrix is of Vandermonde type.

We arrange our points z_1, \dots, z_N into $n+1$ parts as follows. We first collect all points z_j with $|z_j| \leq 1 - \epsilon$, relabel these points $z_{0,0}, \dots, z_{0,m_0}$. If $|z_j| > 1 - \epsilon$, fix any $l \in \{1, \dots, n\}$ with $z_j \in [0, 1]a_l$. We add z_j to the l -th part. So for each $l \in \{1, \dots, n\}$ we obtain points $z_{l,0}, \dots, z_{l,m_l}$ on $[0, 1]a_l$.

Note that $(m_0 + 1) + \dots + (m_n + 1) = N$ and $m_l \geq -1$ for all l . We set $p_l = (m_l + 1)/N$.

Fischer's Inequality, Lemma 2.6, implies

$$v \leq \frac{2}{N(N-1)} \sum_{\substack{l=0 \\ m_l \geq 0}}^n \frac{1}{2} \log |\det A_N(\overline{z_{l,0}}, \dots, \overline{z_{l,m_l}})^t A_N(z_{l,0}, \dots, z_{l,m_l})|$$

where $A_N(z_{l,0}, \dots, z_{l,m_l}) \in \text{Mat}_{N, m_l+1}(\mathbb{C})$ uses notation (2.1).

In this proof, and as usual, the constant implicit in $O(\cdot)$ is absolute.

Recall that $|z_{l,j}| \leq 1$. We apply Proposition 2.1(i) and (ii) to the terms $l \in \{1, \dots, n\}$ and part (iii) to $l = 0$, if $m_l \geq 0$, respectively. Thus

$$(4.2) \quad v \leq p_0 \frac{m_0}{N-1} \log(1 - \epsilon) + \sum_{l=1}^n \min \left\{ 0, p_l \chi(p_l) + \frac{2}{N(N-1)} \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| \right\} + O\left(\frac{\log N}{N}\right),$$

where we used Lemma 3.4 to bound $\gamma_{N,m}$ from above; a term coming from some l with $m_l \leq -1$ can be omitted; we used $(m_0 + 1) + \dots + (m_n + 1) = N$ to bound the error term.

Let us treat the terms on the right-hand side separately.

For $l = 0$ and if $m_0 \geq 0$ we use $\frac{m_0}{N-1} = p_0 - \frac{N-(m+1)}{N(N-1)}$ and $\log(1 - \epsilon) \leq -\epsilon$ to find

$$(4.3) \quad p_0 \frac{m_0}{N-1} \log(1 - \epsilon) \leq -\epsilon p_0^2 + \epsilon p_0 \frac{N - (m+1)}{N(N-1)} \leq -\epsilon p_0^2 + O\left(\frac{1}{N}\right).$$

Let $l \in \{1, \dots, n\}$ with $m_l \geq 0$. The points $z_{l,0}, \dots, z_{l,m_l}$ lie on a line segment of length ϵ . By (4.1) we find

$$\sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| \leq \frac{m_l(m_l + 1)}{2} \log \frac{\epsilon}{4} + O((m_l + 1) \log(m_l + 1)).$$

Recall $\frac{m_l}{N-1} = p_l - \frac{N-(m+1)}{N(N-1)}$ and $m_l + 1 \leq N$. So we get

$$(4.4) \quad \frac{2}{N(N-1)} \sum_{0 \leq i < j \leq m_l} \log |z_{l,j} - z_{l,i}| \leq p_l^2 \log \frac{\epsilon}{4} + O\left(p_l \frac{\log(4N/\epsilon)}{N}\right).$$

We plug (4.3) and (4.4) into (4.2) and find

$$v \leq -\epsilon p_0^2 + \sum_{l=1}^n p_l \min \left\{ 0, \chi(p_l) + p_l \log \frac{\epsilon}{4} \right\} + O\left(\frac{\log(4N/\epsilon)}{N}\right)$$

as $p_1 + \dots + p_n \leq 1$.

Next recall that $x \mapsto \chi(x)$ is concave on $[0, 1]$ as noted below (3.1). Therefore, so is $x \mapsto \chi(x) + x \log \frac{\epsilon}{4}$ and $x \mapsto \min\{0, \chi(x) + x \log \frac{\epsilon}{4}\}$. Jensen's Inequality implies

$$v \leq -\epsilon(1 - p)^2 + p \min \left\{ 0, \chi\left(\frac{\sigma^2}{p}\right) + \frac{\sigma^2}{p} \log \frac{\epsilon}{4} \right\} + O\left(\frac{\log(4N/\epsilon)}{N}\right)$$

with $p = p_1 + \dots + p_n = 1 - p_0$ and $\sigma^2 = p_1^2 + \dots + p_n^2$; if $p = 0$ then the bound holds when omitted the term $p \min\{\dots\}$ here and corresponding terms below.

By (3.2) we have $\chi(x) \leq -x \log(4x/e^{3/2})$ for all $x \in (0, 1]$. So

$$v \leq -\epsilon(1 - p)^2 - \sigma^2 \log^+ \left(\frac{16\sigma^2}{e^{3/2}p\epsilon} \right) + O\left(\frac{\log(4N/\epsilon)}{N}\right).$$

The Cauchy-Schwarz Inequality implies $\sigma^2 \geq p^2/n$ and thus

$$v \leq -\epsilon(1 - p)^2 - \frac{p^2}{n} \log^+ \left(\frac{16p}{e^{3/2}\epsilon n} \right) + O\left(\frac{\log(4N/\epsilon)}{N}\right).$$

We recall $\epsilon = 1/n$. So

$$(4.5) \quad v \leq -\frac{1}{n} \left((1 - p)^2 + p^2 \log^+ \left(\frac{16p}{e^{3/2}} \right) \right) + O\left(\frac{\log(nN)}{N}\right).$$

If $p < e^{3/2}/16 = 0.2801\dots$, then

$$(4.6) \quad v \leq -\frac{1}{n} \left(1 - \frac{e^{3/2}}{16}\right)^2 + O\left(\frac{\log(nN)}{N}\right) \leq -\frac{1}{2n} + O\left(\frac{\log(nN)}{N}\right).$$

For example, if $p = 0$, then $v \leq -1/n + O(\log(nN)/N)$.

If $p \geq e^{3/2}/16$, then we can replace \log^+ by \log in (4.5) and conclude

$$v \leq -\frac{f(p)}{n} + O\left(\frac{\log(nN)}{N}\right) \quad \text{where} \quad f(p) = (1-p)^2 + p^2 \log\left(\frac{16p}{e^{3/2}}\right).$$

The second derivative of $x \mapsto f(x)$ is $\log(2^8 e^2 x^2)$. So f is convex $(0.1, \infty)$. As $f'(0.487) < 0 < f'(0.488)$ the derivative f' has a zero $p_0 \in [0.487, 0.488]$. Thus $f(p) \geq f(p_0)$. Using that $x \mapsto x^2 \log(16xe^{-3/2})$ is increasing on $(e/16, \infty) \supseteq (0.2, \infty)$ we obtain $f(p) \geq f(p_0) \geq (1 - 0.488)^2 + 0.487^2 \log(16 \cdot 0.487e^{-3/2}) > 0.39$. So

$$(4.7) \quad v \leq -\frac{0.39}{n} + O\left(\frac{\log(nN)}{N}\right).$$

Regardless of the size of p we have (4.7) by (4.6).

As $z_1, \dots, z_N \in \mathcal{S}$ are pairwise distinct, but otherwise arbitrary, we conclude

$$\log d_N(\mathcal{S}) = \sup_{z_1, \dots, z_N \in \mathcal{K}} \frac{2}{N(N-1)} \log \prod_{1 \leq i < j \leq N} |z_j - z_i| \leq -\frac{0.39}{n} + O\left(\frac{\log(nN)}{N}\right).$$

Taking the limit $N \rightarrow \infty$ yields $\log d(\mathcal{S}) \leq -0.39/n$, as desired. \square

5 - Algebraic Numbers of Small Height

We conclude this paper by making some remarks on algebraic numbers of small height. Consider a sequence of algebraic numbers of absolute logarithmic Weil height tending to 0 without an infinite constant subsequence. Bilu's Theorem [Bil97] implies that the complex Galois orbits equidistribute towards the Haar measure on the unit circle. In this section we explore consequences of our estimates for algebraic numbers of small height.

Let α be an algebraic number. There is a unique irreducible element P of $\mathbb{Z}[X]$ that vanishes at α and has positive leading term. The absolute logarithmic Weil height $h(\alpha)$, or just height, of α is $m(P)/\deg P$; recall that $m(P)$ is the Mahler measure of P .

We will abbreviate $N = \deg P$ and assume $N \geq 2$.

In what follows we will think of $h(\alpha)$ as being small.

Let $\epsilon \in (0, 1]$ and $n \in \mathbb{N}$. For each $l \in \{1, \dots, n\}$ let \mathcal{D}_l be a closed disk in the complex plane of radius at most ϵ . We will assume that all complex roots of P lie in $\bigcup_{l=1}^n \mathcal{D}_l$.

We fix for each complex root z of P an $l \in \{1, \dots, n\}$ with $z \in \mathcal{D}_l$; this l may not be unique. Let $m_l + 1$ be the number of such roots assigned to l , so $m_l \geq -1$. We also set $p_l = (m_l + 1)/N$ and observe $p_1 + \dots + p_n = 1$.

This assignment induces a partition as in Corollary 1.3 with $\epsilon_1 = \dots = \epsilon_n = \epsilon$. Note that $p = 1$ and $|\text{disc}(P)| \geq 1$ as $P \in \mathbb{Z}[X]$ is squarefree. In the conclusion of Corollary 1.3 we divide by $N(N-1)/2$ and find

$$0 \leq \chi(\sigma^2) + \sigma^2 \log \epsilon + 2h(\alpha) + O\left(\frac{\log(N/\epsilon)}{N}\right)$$

where $\sigma = (p_1^2 + \dots + p_n^2)^{1/2}$.

We use $\chi(x) \leq -x \log x + (3/2 - \log 4)x$, a consequence of (3.2), rearrange, and deduce

$$\sigma^2 \log \left(\frac{4\sigma^2}{e^{3/2}\epsilon} \right) \leq 2h(\alpha) + O\left(\frac{\log(N/\epsilon)}{N}\right).$$

The Cauchy–Schwarz Inequality implies $\sigma^2 \geq 1/n$. Thus

$$(5.1) \quad \sigma^2 \log \left(\frac{4}{e^{3/2}n\epsilon} \right) \leq 2h(\alpha) + O\left(\frac{\log(N/\epsilon)}{N}\right).$$

Let us assume that $n\epsilon < 4e^{-3/2} = 0.8925206\dots$ and suppose $\kappa > 0$ satisfies $n\epsilon \leq 4e^{-3/2}e^{-\kappa}$. Thus $\sigma^2\kappa \leq 2h(\alpha) + O\left(\frac{\log(N/\epsilon)}{N}\right)$. We have $\sigma^2 = \sum_{l=1}^n p_l^2 = \frac{1}{n} + \sum_{l=1}^n (p_l - \frac{1}{n})^2$ and therefore

$$(5.2) \quad \frac{1}{n} + \sum_{l=1}^n \left(p_l - \frac{1}{n} \right)^2 \leq \frac{2}{\kappa} h(\alpha) + O\left(\frac{\log(N/\epsilon)}{\kappa N}\right).$$

Consider a sequence of algebraic numbers α with $h(\alpha) \rightarrow 0$ and $N = [\mathbb{Q}(\alpha) : \mathbb{Q}] \rightarrow \infty$. Suppose that all complex Galois conjugates of α lie in $\bigcup_{l=1}^n \mathcal{D}_l$. We allow n and ϵ to vary along this sequence subject to $n\epsilon \leq 4e^{-3/2}e^{-\kappa} < 1$ with fixed $\kappa > 0$ while also assuming that $\log(N/\epsilon)/N \rightarrow 0$ as $N \rightarrow \infty$. Then we conclude two things from (5.2). First, $n \rightarrow \infty$, *i.e.*, the number of disks of radius $\leq \epsilon$ required to cover all Galois conjugates of α tends to ∞ . Second, the normalized variance $\sum_{l=1}^n (p_l - \frac{1}{n})^2$ tends to 0. This means that each disk \mathcal{D}_l gets its fair share of Galois conjugates on average.

Bombieri and Zannier obtained a non-archimedean result in the spirit of (5.2), see Theorem 3 [BZ01]. They used their height inequality to exhibit

fields of algebraic numbers that do not contain any elements of sufficiently small positive height.

Let S be a non-empty and bounded subset of \mathbb{C} . For $\epsilon > 0$ let $n(S, \epsilon)$ denote the minimal number of closed disks of radius ϵ needed to cover S . The upper box dimension of S is

$$\overline{\dim}_{\text{box}}(S) = \limsup_{\epsilon \rightarrow 0^+} \frac{\log n(S, \epsilon)}{-\log \epsilon}.$$

It makes no difference if one takes disks of radius ϵ or boxes of side length ϵ .

For any subset $S \subseteq \mathbb{C}$ let $S(\overline{\mathbb{Q}})$ denote the set of algebraic numbers in \mathbb{C} whose Galois conjugates all lie in S .

Theorem 5.1. *Let S be a non-empty and bounded subset of \mathbb{C} with $\overline{\dim}_{\text{box}}(S) < 1$. There exists $\delta = \delta(S) > 0$ such that if $\alpha \in S(\overline{\mathbb{Q}})$ then $h(\alpha) \geq \delta$ or $h(\alpha) = 0$. Moreover, the second case occurs only finitely often.*

Proof. By hypothesis there exists $d \in (0, 1)$ with $n(S, \epsilon) \leq \epsilon^{-d}$ for all sufficiently small $\epsilon > 0$. So S is covered by $n = n(S, \epsilon) \leq \epsilon^{-d}$ closed disks \mathcal{D}_l of radius ϵ . We will fix such an ϵ soon.

Let $\alpha \in S(\overline{\mathbb{Q}})$ with $N = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. If $N = 1$, then $h(\alpha) \geq \log 2$ or $h(\alpha) = 0$; the latter case only happens for $\alpha \in \{0, \pm 1\}$. So we may assume $N \geq 2$.

We use the notation introduced around (5.1) and apply this bound. In our case it implies $\sigma^2 \log(4e^{-3/2}\epsilon^{d-1}) \leq 2h(\alpha) + O(\log(N/\epsilon)/N)$.

We now fix $\epsilon \in (0, 1]$ small enough to ensure $4e^{-3/2}\epsilon^{d-1} \geq e^2$, this is possible as $d < 1$. Therefore, $\sigma^2 \leq h(\alpha) + O(\log(N/\epsilon)/N)$. The Cauchy–Schwarz Inequality implies $\sigma^2 \geq 1/n \geq \epsilon^d \geq \epsilon$. So $\epsilon \leq h(\alpha) + O(\log(N/\epsilon)/N)$.

Having fixed ϵ we find $h(\alpha) \geq \epsilon/2$ for all sufficiently large N . So we in the first case of the theorem. If $h(\alpha) < \epsilon/2$, then $N = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ is bounded from above and we may apply Northcott's Theorem which states that a set of algebraic numbers of bounded height and degree is finite. So $h(\alpha)$ cannot be an arbitrarily small positive real number and it is 0 for at most finitely many α finitely often. \square

We conclude by making some remarks on the previous theorem.

The theorem above does not hold when replacing the upper box dimension by the Hausdorff dimension. Indeed, the group of roots of unity in \mathbb{C} is a countable infinite set and therefore has Hausdorff dimension 0. However, its elements have height 0.

There is a bounded set of upper box dimension strictly less than 1 that contains infinitely many rational numbers. Indeed, the Cantor set consists of real numbers in $[0, 1]$ which have a ternary expansion omitting 1. The Cantor

set contains all positive powers of $1/3$ and is known to have upper box counting dimension $(\log 2)/(\log 3) < 1$.

There is a bounded set of upper box dimension strictly less than 1 that contains all Galois conjugates of infinitely many algebraic integers of bounded height. Indeed, for all complex c of sufficiently large modulus, the filled Julia set K_f of $f: z \mapsto z^2 + c$ is a compact set of Hausdorff dimension < 1 . Moreover, its Hausdorff dimension equals its upper box counting dimension if $|c|$ is large enough. We fix such a $c \in \mathbb{Z}$. It is well-known that f admits infinitely many preperiodic points in the ring of all algebraic integers. All preperiodic points of f lie in K_f and so do their Galois conjugates. Finally, the Call–Silverman height vanishes on all preperiodic points and differs from the height by a bounded function. So the preperiodic points in question have bounded height.

Acknowledgments. A first iteration of this work was presented at a virtual reading course on Dimitrov’s proof of the Schinzel–Zassenhaus Conjecture. The course began at the onset of the current pandemic and the author thanks all participants for creating a positive environment despite the circumstances. He also thanks Yann Bugeaud for comments and references and the referee for providing the reference [Sch18].

References

- [BBG10] V. BERESNEVICH, V. BERNIK and F. GÖTZE, *The distribution of close conjugate algebraic numbers*, Compos. Math. **146** (2010), 1165–1179.
- [BD11] Y. BUGEAUD and A. DUJELLA, *Root separation for irreducible integer polynomials*, Bull. Lond. Math. Soc. **43** (2011), 1239–1244.
- [BD14] Y. BUGEAUD and A. DUJELLA, *Root separation for reducible integer polynomials*, Acta Arith. **162** (2014), 393–403.
- [Bil97] Y. BILU, *Limit distribution of small points on algebraic tori*, Duke Math. J. **89** (1997), 465–476.
- [BM04] Y. BUGEAUD and M. MIGNOTTE, *On the distance between roots of integer polynomials*, Proc. Edinb. Math. Soc. (2) **47** (2004), 553–556.
- [BM10] Y. BUGEAUD and M. MIGNOTTE, *Polynomial root separation*, Int. J. Number Theory **6** (2010), 587–602.
- [BZ01] E. BOMBIERI and U. ZANNIER, *A note on heights in certain infinite extensions of \mathbb{Q}* , Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **12** (2001), 5–14.
- [Dim] V. DIMITROV, *A proof of the Schinzel-Zassenhaus conjecture on polynomials*, arXiv:1912.12545, preprint, 2019.

- [Dub84] V. N. DUBININ, *Change of harmonic measure in symmetrization* (Russian), Mat. Sb. (N.S.) **124(166)** (1984), 272–279.
- [EG17] J.-H. EVERTSE and K. GYÖRY, *Discriminant equations in Diophantine number theory*, New Math. Monogr., **32**, Cambridge University Press, Cambridge, 2017.
- [Eve04] J.-H. EVERTSE, *Distances between the conjugates of an algebraic number*, Publ. Math. Debrecen **65** (2004), 323–340.
- [Fra79] J. S. FRAME, *The Hankel power sum matrix inverse and the Bernoulli continued fraction*, Math. Comp. **33** (1979), 815–826.
- [Güt67] R. GÜTING, *Polynomials with multiple zeros*, Mathematika **14** (1967), 181–196.
- [Hil94] D. HILBERT, *Ein Beitrag zur Theorie des Legendre’schen Polynoms*, Acta Math. **18** (1894), 155–159.
- [Kra99] C. KRATTENTHALER, *Advanced determinant calculus*, The Andrews Festschrift (Maratea, 1998), Sémin. Lothar. Combin. **42** (1999), Art. B42q, 67 pp.
- [Mah64] K. MAHLER, *An inequality for the discriminant of a polynomial*, Michigan Math. J. **11** (1964), 257–262.
- [Mig95] M. MIGNOTTE, *On the distance between the roots of a polynomial*, Appl. Algebra Engrg. Comm. Comput. **6** (1995), 327–332.
- [Mir55] L. MIRSKY, *An introduction to linear algebra*, Oxford, at the Clarendon Press, 1955.
- [Pro89] R. A. PROCTOR, *Equivalence of the combinatorial and the classical definitions of Schur functions*, J. Combin. Theory Ser. A **51** (1989), 135–137.
- [Ran95] T. RANSFORD, *Potential theory in the complex plane*, London Math. Soc. Stud. Texts, **28**, Cambridge University Press, Cambridge, 1995.
- [Rum79] S. M. RUMP, *Polynomial minimum root separation*, Math. Comp. **33** (1979), 327–336.
- [Sch18] I. SCHUR, *Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, Math. Z. **1** (1918), 377–402.
- [Sch06] A. SCHÖNHAGE, *Polynomial root separation examples*, J. Symbolic Comput. **41** (2006), 1080–1090.

PHILIPP HABEGGER
Department of Mathematics and Computer Science
University of Basel
Spiegelgasse 1
4051 Basel, Switzerland
e-mail: philipp.habegger@unibas.ch