Christos Philippou

A proof of the Schinzel-Zassenhaus conjecture

Master Thesis
July 23, 2024

Supervisors:

Dr. J.H. Evertse
Dr. D.T.B.G. Lilienfeldt

Leiden University
Mathematical Institute

Contents

1 Introduction

In this thesis, we give a proof of the Schinzel-Zassenhaus conjecture by combining the recent strategy of Dimitrov [4] with a transfinite diameter bound due to Habegger [9]. While the result obtained in this way is weaker than the one of Dimitrov, our proof is self-contained and only relies on elementary methods.

1.1 Main result

Let α be a non-zero algebraic integer, not a root of unity and let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_d$ denote the conjugates of α. In 1965, Schinzel and Zassenhaus [20] conjectured that

$$\lceil\overline{\alpha}\rceil := \max_{1 \leq i \leq d} |\alpha_i| > 1 + \frac{c}{d} \tag{1.1}$$

for some absolute constant $c > 0$. In simple words this says that every non-cyclotomic monic polynomial with integer coefficients should have at least one root that is outside the unit circle by a distance equal to a constant number divided by the degree of the polynomial.

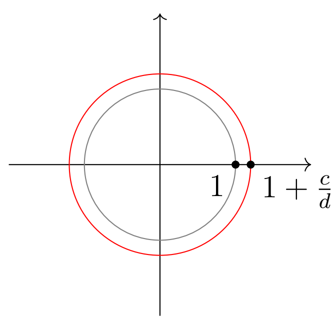

Figure 1: Every non-cyclotomic monic polynomial of degree d with integer coefficients should have at least one root that is outside the outer circle.

Any non-zero complex number z can be written in polar form $z = re^{i\theta}$, where $r > 0$ is the modulus and $\theta \in \mathbb{R}$ is called the argument of z, denoted by arg (z) (defined uniquely up to addition of an integer multiple of 2π). The main result of this thesis is the following bound:

Theorem 1.1. *Let α be a non-zero algebraic integer of degree d, not a root of unity, whose conjugates have d' different arguments between them. Then*

$$\overline{|\alpha|} \geq e^{0.39/8d'}.$$

As an immediate consequence, we deduce the following:

Corollary 1.2. *The Schinzel-Zassenhaus conjecture (1.3) holds with $c = \frac{0.39}{8}$.*

Proof. From $d' \leq d$, we have $\overline{|\alpha|} \geq e^{0.39/8d}$. We now want to show that $e^{0.39/8d} \geq 1 + \frac{0.39}{8d} = 1 + \log e^{0.39/(8d)}$. Using that $a \geq 1 + \log a$ for $a \geq 1$, we get the result. □

A non-zero algebraic integer α is called reciprocal if α^{-1} is a conjugate of α.

Corollary 1.3. *If α is a reciprocal algebraic integer of degree $d > 1$ with $d'' > 0$ conjugates not on the unit circle, then*

$$\overline{|\alpha|} \geq e^{0.39/(8d-4d'')}.$$

In particular, if $d'' = d$, then $\overline{|\alpha|} \geq e^{0.39/4d}$.

Proof. First note that the complex conjugate of an algebraic integer is one of its conjugates and that because α is a reciprocal algebraic integer then by definition α^{-1} is also a conjugate of α.
Let $\alpha = a + ib$. Its inverse is $\alpha^{-1} = \frac{1}{a+ib} = \frac{a-ib}{a^2+b^2}$ and its complex conjugate is $\overline{\alpha} = a - ib$. So we have that

$$\arg(\alpha^{-1}) = \arg\left(\frac{1}{a^2+b^2}\right) + \arg(a-ib) = 0 + \arg(\overline{\alpha}) = \arg(\overline{\alpha})$$

When α is on the unit circle we have that $\overline{\alpha} = \alpha^{-1}$ and thus when counting the different arguments we count one for each of the conjugates. But when we count the different arguments of the conjugates that are not on the unit circle then $\overline{\alpha} \neq \alpha^{-1}$ so for every one of these d'' conjugates there is another one with the same argument. Therefore the conjugates of α have at most $d - \frac{1}{2}d''$ different arguments. □

1.2 History of the problem

In 1857, Kronecker [12] proved that if α is a non-zero algebraic integer then

$$\overline{|\alpha|} = 1 \iff M(a) = 1 \iff \alpha \text{ is a root of unity,}$$

where $M(\alpha) := \prod_{k=1}^{d} \max\{1, |\alpha_k|\}$ is the Mahler measure of an algebraic integer α with conjugates $\alpha_1, \ldots, \alpha_d$.

Suppose now that α is not a root of unity. In 1933, Lehmer [14] asked if it is true that

$$M(\alpha) \geq 1 + C, \tag{1.2}$$

where $C > 0$ is an absolute constant. Inequality (1.2) is known as Lehmer's conjecture. In 1965, Schinzel and Zassenhaus [20] proved that if α is a non-zero algebraic integer, not a root of unity, and if $2s$ of its conjugates are complex, then

$$\overline{|\alpha|} > 1 + 4^{-s-2}.$$

They further stated the following: '...we cannot disprove the inequality

$$\max_{1 \leq i \leq d} |\alpha_i| > 1 + \frac{c}{d}, \tag{1.3}$$

for some absolute constant $c > 0$'. This statement is known as the Schinzel-Zassenhaus conjecture.

Note that Lehmer's conjecture (1.2) is stronger than the Schinzel-Zassenhaus conjecture (1.3). To show this, suppose that α satisfies $M(\alpha) \geq C > 1$. Note that $\overline{|\alpha|} \leq 1 + \frac{\log C}{d}$ implies $M(\alpha) < \left(1 + \frac{\log C}{d}\right)^d < C$. Thus, the Schinzel–Zassenhaus conjecture holds with $c = \log(C)$. Consequently, results in the direction of Lehmer's conjecture give corresponding results for the Schinzel–Zassenhaus conjecture.

For α a non-reciprocal algebraic integer, Smyth [21] showed in 1971 that

$$M(\alpha) \geq \theta,$$

where $\theta = 1.324\ldots$ denotes the real zero of the polynomial $z^3 - z - 1$ and that this θ is optimal. Hence,

$$\overline{|\alpha|} \geq M(\alpha)^{1/d} > 1 + \frac{\log \theta}{d}.$$

In 1978, Dobrowolski [5] proved that if a non-zero algebraic integer α of degree d satisfies

$$\overline{|\alpha|} \leq 1 + \frac{1}{4ed^2}, \tag{1.4}$$

then α is a root of unity. The proof is short so we include it here.

Proof. Let $s_k = \alpha_1^k + \cdots + \alpha_d^k$ where $\alpha = \alpha_1, \ldots, \alpha_d$ are the conjugates of α. Let p be a prime between $2ed$ and $4ed$. Recall that $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$. If (1.4) holds, then for $k \leq d$ we have

$$|s_k| \leq d\Big(1 + \frac{1}{4ed^2}\Big)^d < d\Big(1 + \frac{1}{d}\Big)^d < de$$

and

$$|s_{kp}| \leq d\Big(1 + \frac{1}{4ed^2}\Big)^{4ed^2} < de.$$

Hence

$$|s_{kp} - s_k| \leq 2ed < p.$$

By Theorem A.15, $s_{kp} \equiv s_k \pmod{p}$, so that $s_{kp} = s_k$ for $k = 1, \ldots d$. Since these values determine the coefficients of the minimal polynomials of α and α^p, using Theorem A.14 we see that they have the same minimal polynomial. Hence, α^p is a zero of the minimal polynomial of α. Thus, α is conjugate to α^p and also to α^{p^m} for all m. This holds because if $L = \mathbb{Q}(\alpha_1, \ldots, \alpha_d)$, then for every $1 \leq i, j \leq d$ there is a σ in the Galois group such that $\sigma(\alpha_i) = \alpha_j$. Thus there is a σ such that $\sigma(\alpha) = \alpha^p$, $\sigma(\alpha^p) = \alpha^{2p}$ and so on. So two of these must be equal and therefore α must be a root of unity. □

In 1979, Dobrowolski [6] proved that for any $\epsilon > 0$ there exist effective constants $d_1(\epsilon)$ and $d_2(\epsilon)$ such that, for $d > d_1(\epsilon)$

$$M(\alpha) > 1 + (1 - \epsilon)\left(\frac{\log\log d}{\log d}\right)^3, \tag{1.5}$$

and, for $d > d_2(\epsilon)$

$$\overline{|\alpha|} > 1 + \frac{2 - \epsilon}{d}\left(\frac{\log\log d}{\log d}\right)^3. \tag{1.6}$$

Later, in 1982, Cantor and Straus [2] replaced the constant $1 - \epsilon$ by $2 - \epsilon$ in (1.5) and $2 - \epsilon$ by $4 - \epsilon$ in (1.6) respectively. In 1993, Dubickas [7] calulated the best such $c' - \epsilon$ in (1.6) with $c' = 64/\pi^2$.

In 2019 Dimitrov [4] obtained the bound

$$\overline{|\alpha|} \geq 2^{1/4d},$$

which implies the Schinzel-Zassenhaus conjecture (1.3) with $c = \frac{1}{4}\log 2$. In his proof, Dimitrov crucially used a bound on the transfinite diameter of a compact subset of $\mathbb{C}$ known as a hedgehog.

1.3 Strategy of the proof

In this thesis, we reproduce Dimitrov's proof of the Schinzel-Zassenhaus conjecture using a slightly weaker transfinite diameter bound due to Habegger [9], whose proof has the virtue of relying only on elementary methods. As a consequence, Theorem 1.1 is slightly weaker than the result of Dimitrov [4]. We will now outline the strategy of proof of Theorem 1.1.

Let α be a non-zero algebraic integer of degree d having minimal polynomial $P(x) = \prod_{i=1}^{d}(x-\alpha_i)$. Define the polynomial $P_m(x) := \prod_{i=1}^{d}(x-\alpha_i^m) \in \mathbb{Z}[x]$ to be the polynomial whose zeros are the mth powers of the conjugates of α. We introduce a bit of terminology. The reciprocal of a polynomial $Q(x) \in \mathbb{C}[x]$ is by definition the polynomial $Q^*(x) := x^{\deg(Q)}Q(1/x) \in \mathbb{C}[x]$. Given complex numbers $a_1, \ldots, a_n$, the associated "hedgehog" $\mathcal{K} = \mathcal{K}(a_1, \ldots, a_n)$ is the compact subset of $\mathbb{C}$ formed by taking the union of the line segments joining 0 to a_j for $j \in \{1, \ldots, n\}$.

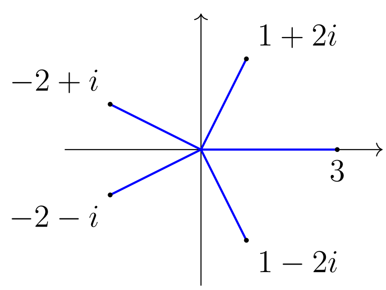

Figure 2: Example of the hedgehog $\mathcal{K}(a_1, \ldots, a_n)$ where the a_i are the roots of the polynomial $x^5 - 5x^4 - 4x^2 + 51x + 45$.

Consider the polynomial $P_2(z)$ where $P(z)$ is the minimal polynomial of α. In the case where $P_2(z)$ is reducible, Theorem 1.1 follows by induction on the degree d. We use that since $P_2(z)$ is reducible, with α^2 one of its zeros, the degree of α^2 is less than that of α and we can therefore apply the induction hypothesis to α^2.

In the case where P_2 is irreducible consider the series

$$F\left(\frac{1}{z}\right) := \sqrt{P_2^*\left(\frac{1}{z}\right) P_4^*\left(\frac{1}{z}\right)} \in \mathbb{C}\left[\left[\frac{1}{z}\right]\right].$$

The rationality of this series is intimately related to the polynomial $P(z)$ being cyclotomic. More precisely, if $F(1/z)$ is rational, then $P_2(z) = P_4(z)$, which implies that $P(z)$ is cyclotomic (Lemma 5.5).

The rationality of the series is also related to the transfinite diameter of the hedgehog $\mathcal{K}(\alpha_1^2, \ldots, \alpha_d^2, \alpha_1^4, \ldots, \alpha_d^4)$. More precisely, the coefficients of $F(1/z)$, which *a priori* belong to $\mathbb{C}$, can be shown to be integral (Lemma 5.4). Also, there exists criteria for the rationality of integral power series in terms of transfinite diameters of compact subsets of $\mathbb{C}$ (Theorem 5.2). In the case of $F(1/z)$, if the transfinite diameter of the hedgehog $\mathcal{K}(\alpha_1^2, \ldots, \alpha_d^2, \alpha_1^4, \ldots, \alpha_d^4)$ is less than 1, then $F(1/z)$ is rational and, by Habegger's result (Theorem 4.4), this transfinite diameter is less than 1 when $\lceil\alpha\rceil < e^{0.39/(8d)}$. Therefore assuming $\lceil\alpha\rceil < e^{0.39/(8d)}$ yields a contradiction because α is not a root of unity, and Theorem 1.1 follows.

1.4 Outline of thesis

In Section 2, we introduce the transfinite diameter of a compact set. We prove some of its basic properties and compute the transfinite diameter in two illustrative examples: the unit circle and the unit segment. Section 3 is dedicated to bounding a number of values that will be key in bounding the transfinite diameter of the hedgehog in Section 4. In Section 5, , we give the rationality criterion for integral series and finish with the proof of Theorem 1.1. The Appendix A contains miscellaneous results used throughout the thesis.

2 Transfinite diameter

In this section we will define what is the transfinite diameter and prove some useful properties. This section is based on chapter 10 of [16].

2.1 Definition and properties

Definition 2.1 (Transfinite diameter)**.** Let E be a compact subset of the complex plane $\mathbb{C}$, symmetric about the real axis. For a given integer n we define G_n as

$$G_n = G_n(E) = \max_{z_1,\ldots,z_n \in E} \left(\prod_{1 \le j < k \le n} |z_j - z_k| \right)^{2/n(n-1)}. \tag{2.1}$$

Then the transfinite diameter $\tau(E)$ of E is defined as

$$\lim_{n \to \infty} G_n.$$

We will show later in Proposition 2.6 that the limit exists.

The transfinte diameter is a way of quantifying the size of compact sets and was introduced by M. Fekete [8]. It appears in analysis and potential theory, where it is also known as the logarithmic capacity, the exterior mapping radius, or the Chebyshev constant of E. The transfinite diameter has numerous applications in number theory. One noteworthy such application is the following. Let E be a compact subset of the complex plane, symmetric about the real axis, and of transfinite diameter less than 1. Then E contains only finitely many conjugate sets of algebraic integers [16].

Some basic properties of the transfinite diameter are:

- for $\lambda, \mu \in \mathbb{C},\ \tau(\lambda E + \mu) = |\lambda| \tau(E)$.
- $\tau(E') \le \tau(E)$ for any compact subset E' of E.

Now we give some more definitions and some less obvious properties of the transfinite diameter.

Definition 2.2 (Chebychev polynomials)**.** For $n \in \mathbb{Z}_{\geq 0}$ let $T_n(z) \in \mathbb{C}[z]$ be monic polynomial of degree n chosen so that its maximum modulus on E

$$m_n := \max_{z \in E} |T_n(z)|$$

is as small as possible. The polynomials $T_n(z)$ are called the Chebychev polynomials of E.

Chebyshev polynomials originated from a problem in classical mechanics, and due to their properties they have various uses in number theory.

For the rest of this section let E be a compact subset of $\mathbb{C}$ and let G_n be as in (2.1).

Definition 2.3. For $z_1, \ldots, z_n \in E$, define

$$V(z_1, \ldots, z_n) = \prod_{\substack{j,k=1 \\ j<k}}^{n} (z_j - z_k) \tag{2.2}$$

and let V_n be the maximum of $|V(z_1, \ldots, z_n)|$ as $z_1, \ldots, z_n$ range over E. Because E is compact, this maximum exists.

Proposition 2.4. *The sequence $\{G_n\}$ is monotonically decreasing.*

Proof. We have

$$G_n = V_n^{\frac{2}{n(n-1)}}.$$

Choose $n+1$ new points $z_1, \ldots, z_n, z_{n+1} \in E$ such that $|V(z_1, \ldots, z_n, z_{n+1})| = V_{n+1}$. Then, since

$$V(z_1, \ldots, z_n, z_{n+1}) = (z_1 - z_2)(z_1 - z_3) \cdots (z_1 - z_{n+1}) V(z_2, \ldots, z_n, z_{n+1}),$$

and $|V(z_2, \ldots, z_n, z_{n+1})| \leq V_n$ we have

$$V_{n+1} \leq |z_1 - z_2| \cdot |z_1 - z_3| \cdots |z_1 - z_{n+1}| V_n.$$

Similarly, doing the same for all sets of n of these $n+1$ points, we have

$$\begin{array}{ccccc} V_{n+1} \leq |z_2 - z_1| & \cdot & |z_2 - z_3| & \cdots & |z_2 - z_{n+1}| V_n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{n+1} \leq |z_{n+1} - z_1| & \cdot & |z_{n+1} - z_2| & \cdots & |z_{n+1} - z_n| V_n. \end{array}$$

Multiplying all these inequalities we obtain

$$V_{n+1}^{n+1} \leq V_{n+1}^{2} V_{n}^{n+1}.$$

Hence $V_{n+1}^{n-1} \leq V_n^{n+1}$, giving

$$G_{n+1} = V_{n+1}^{\frac{2(n-1)}{(n-1)(n+1)n}} \leq V_n^{\frac{2(n+1)}{(n-1)(n+1)n}} = V_n^{\frac{2}{(n-1)n}} = G_n.$$

□

Lemma 2.5. *Let $\{a_n\}_{n\geq 1}$ be a sequence of positive numbers with finite limit $a > 0$. Then the sequence $(a_1^1 a_2^2 \cdots a_n^n)^{\frac{2}{n)(n+1)}}$ converges to a as $n \to \infty$.*

Proof. Without loss of generality because of the 2 in the exponent we can assume that $a_i > 1$ for all $i \geq 1$. Now let $c_i = \log(a_i) > 0$ and $c = \log(a)$. We want to prove that if $c_i \to c$ then $\frac{1}{n(n+1)} \sum_{i=1}^{n} i \cdot c_i \to c$ as $n \to \infty$.
Let $\epsilon > 0$ and let N be such that $|c_n - c| < \epsilon$ for every $n \geq N$.
We have, for $n > N$

$$\begin{aligned}
\left| \left(\frac{2}{n(n+1)} \sum_{i=1}^{n} i c_i \right) - c \right| &= \left| \frac{2}{n(n+1)} \sum_{i=1}^{n} i c_i - \frac{2}{n(n+1)} \sum_{i=1}^{n} i c \right| \\
&= \left| \frac{2}{n(n+1)} \sum_{i=1}^{n} i(c_i - c) \right| \\
&\leq \left| \frac{2}{n(n+1)} \sum_{i=1}^{N} i(c_i - c) \right| + \left| \frac{2}{n(n+1)} \sum_{i=N+1}^{n} i(c_i - c) \right| \\
&\leq \frac{2}{n(n+1)} C + \frac{2}{n(n+1)} \left(\sum_{i=N+1}^{n} i \right) \epsilon \qquad (\ C \text{ a constant}) \\
&< \frac{C}{n(n+1)} + \epsilon < 2\epsilon,
\end{aligned}$$

provided we choose N sufficiently large such that $\frac{C}{n(n+1)} < \frac{C}{N(N+1)} < \epsilon$. □

Proposition 2.6. *Let m_n be as in Definition 2.2. Then ${m_n}^{1/n}$ converges as $n \to \infty$, and the limit is the transfinite diameter $\tau(E)$.*

Proof. We first remark that the sequence $\tau_n := m_n^{1/n}$ is bounded since for the polynomial $(z-z_0)^n$ with $z_0 \in E$ we have

$$\tau_n = m_n^{1/n} \leq (\max_{z\in E} |z-z_0|^n)^{1/n} = \max_{z\in E} |z-z_0| \leq \max_{z,z_0\in E} |z-z_0| = D,$$

where D is the (standard) diameter of E.
Next we show that the sequence converges. Define $a := \liminf_{n\to\infty} \tau_n$ and $b := \limsup_{n\to\infty} \tau_n$. We have $a \leq b$ and want to show that $b \leq a$. Let $\frac{\epsilon}{2} > 0$ and choose n such that $\tau_n < a + \frac{\epsilon}{2}$. Then because $\tau_n < a + \frac{\epsilon}{2}$ we have

$$\begin{aligned} \tau_n = m_n^{1/n} &< a + \frac{\epsilon}{2} \\ \max_{z\in E} |T_n(z)|^{1/n} &< a + \frac{\epsilon}{2} \\ |T_n(z)| &< \left(a + \frac{\epsilon}{2}\right)^n, \text{ for all } z \in E. \end{aligned}$$

If k and l are positive integers and $z_0 \in E$ is fixed, then for all $z \in E$

$$|(z-z_0)^l T_n(z)^k| \leq D^l \left(a + \frac{\epsilon}{2}\right)^{nk}.$$

Hence

$$\begin{aligned} \tau_{nk+l} &= \max_{z\in E} |T_{nk+l}(z)|^{1/(nk+l)} \\ &\leq \max_{z\in E} |(z-z_0)^l T_{nk}(z)|^{1/(nk+l)} \\ &\leq \max_{z\in E} |(z-z_0)^l|^{1/(nk+l)} \cdot \max_{z\in E} |T_{nk}(z)|^{1/(nk+l)} \\ &\leq D^{\frac{l}{nk+l}} \left(a + \frac{\epsilon}{2}\right)^{\frac{nk}{nk+l}}. \end{aligned}$$

Now choosing a subsequence $\tau_{n_\nu} \to b$ as $n_\nu \to \infty$ and putting $n_\nu = nk_\nu + l_\nu$, where $0 < l_\nu \leq n$, we see that for $n_\nu \gg 0$ such that $\tau > b - \frac{\epsilon}{2}$ or $b < t_{n_\nu} + \frac{\epsilon}{2}$

$$\begin{aligned} b &\leq \tau_{nk_\nu + l_\nu} + \frac{\epsilon}{2} \\ &\leq D^{\frac{l}{nk+l}} \left(a + \frac{\epsilon}{2}\right)^{\frac{nk}{nk+l}} + \frac{\epsilon}{2} \\ &\leq D^{\frac{l}{nk+l}} \left(a + \frac{\epsilon}{2}\right)^{\frac{nk}{nk+l}} + \frac{\epsilon}{2} \\ &\leq D^0 \left(a + \frac{\epsilon}{2}\right) + \frac{\epsilon}{2} \qquad \text{as } k_\nu \to \infty \\ &= (a + \epsilon). \end{aligned}$$

Since ϵ can be arbitrarily small, we have $b \leq a$, as was our aim. Hence $a = b$, so that the sequence $\{\tau_\nu\}$ converges.

We must now show that in fact $\{\tau_\nu\}$ converges to $\tau(E)$. This is clearly true if E is finite because for $n > |E|$ there will be terms of the form $|z_i - z_i| = 0$ for $z_i \in E$ in the product $\prod_{1 \leq j < k \leq n} |z_j - z_k|$ and $\tau(E)$ will be equal to 0. So we assume that E is infinite, and V_n is never 0. We claim that

$$m_n \leq \frac{V_{n+1}}{V_n} \leq (n+1)m_n. \tag{2.3}$$

To prove these inequalities, note that

$$m_n \leq \max_{z \in E} |(z - z_1) \cdots (z - z_n)| = \frac{\max_{z \in E} |V(z, z_1, \ldots, z_n)|}{V(z_1, \ldots, z_n)} \leq \frac{V_{n+1}}{V_n}$$

when $z_1, \ldots, z_n \in E$ are chosen so that $|V(z_1, \ldots, z_n)| = V_n$. This proves the first inequality of (2.3). For the second inequality, we note that $V(z_1, \ldots z_{n+1})$ is the determinantant of a Vandermonde matrix defined in Proposition A.3,

$$V(z_1, \ldots z_{n+1}) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ z_1^n & z_2^n & \ldots & z_{n+1}^n \end{vmatrix}$$

and then add suitable multiples of the first n rows to the bottom row to obtain

$$V(z_1, \ldots z_{n+1}) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ T_n(z_1) & T_n(z_2) & \ldots & T_n(z_{n+1}) \end{vmatrix}.$$

On expanding this determinant by its bottom row, we have that

$$|V(z_1, \ldots z_{n+1})| \leq |T_n(z_1)| \cdot |V(z_2, \ldots z_{n+1})| + |T_n(z_2)| \cdot |V(z_1, z_3, \ldots z_{n+1})| \\ + \cdots + |T_n(z_{n+1})| \cdot |V(z_1, \ldots z_n)|.$$

Now, choosing $z_1 \ldots, z_{n+1}$ so that $V(z_1, \ldots z_{n+1}) = V_{n+1}$, we obtain the bound $V_{n+1} \leq (n+1)m_n V_n$, giving the second inequality of (2.3). Thus, we have

$$\tau_n^n \leq \frac{V_{n+1}}{V_n} \leq (n+1)\tau_n^n.$$

Multiplying these inequalities for $n = 2$ up to n', and then replacing n' by n, we obtain

$$(\tau_2^2 \cdots \tau_n^n) \leq \frac{V_3 \cdot V_4 \cdots V_{n+1}}{V_2 \cdot V_3 \cdots V_n} \leq ((n+1)!)(\tau_2^2 \cdots \tau_n^n).$$

We now multiply by V_2

$$(\tau_2^2 \cdots \tau_n^n) \cdot V_2 \leq \ V_{n+1} \leq ((n+1)!)(\tau_2^2 \cdots \tau_n^n) \cdot V_2,$$

and take the $\frac{2}{n(n-1)}$'th power and recall that $G_n = V_n^{\frac{2}{n(n-1)}}$ to obtain

$$(\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}} \cdot V_2^{\frac{2}{n(n+1)}} \leq G_{n+1} \leq ((n+1)!)^{\frac{2}{n(n+1)}}(\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}} \cdot V_2^{\frac{2}{n(n+1)}}.$$

Observe that both $V_2^{\frac{2}{n(n+1)}}$ and $((n+1)!)^{\frac{2}{n(n+1)}}$ tend to 1 as $n \to \infty$.
Now the inequality becomes

$$\lim_{n\to\infty} (\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}} \leq \tau \leq \lim_{n\to\infty} (\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}}$$

and thus $\lim_{n\to\infty}(\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}} = \tau$. From Lemma 2.5 we know that if $\lim_{n\to\infty} \tau_n$ exists then it is equal to the $\lim_{n\to\infty}(\tau_2^2 \cdots \tau_n^n)^{\frac{2}{n(n+1)}}$. Therefore $\lim_{n\to\infty} \tau_n = \lim_{n\to\infty} m_n^{1/n} = \tau$. □

Proposition 2.7. *For any monic polynomial $h(z)$ of degree d, the transfinite diameter of $h^{-1}(E)$ is $\tau(E)^{1/d}$.*

Proof. Let $T_n(z)$ and $T_n^*(z)$ denote the n-th Chebyshev polynomial for E and $E^* := h^{-1}(E)$, respectively, and put

$$m_n^* := \max_{z\in E^*} |T_n^*(z)|.$$

Then

$$m_{dn}^* = \max_{z\in E^*} |T_{dn}^*(z)| \leq \max_{z\in E^*} |T_n(h(z))| = \max_{z\in E} |T_n(z)| = m_n.$$

The second equality holds because the sets $h(E^*) = h\left(h^{-1}(E)\right)$ and E are equal by surjectivity of h. This gives

$$(m_{dn}^*)^{1/dn} \leq (m_n^{1/n})^{1/d}$$

which, on letting $n \to \infty$, shows that $\tau(E^*) \le \tau(E)^{1/d}$.

We must now prove this inequality in the other direction. For an arbitrary fixed $w_0 \in E$, we define $z_1, \ldots, z_d$ by $h(z) - w_0 = \prod_{j=1}^{d}(z - z_j)$ and define $z_1^*, \ldots, z_n^*$ in E^* by $T^*(z) := \prod_{k=1}^{n}(z - z_k^*)$. Then from the identity

$$\left|\prod_{j=1}^{d}\prod_{k=1}^{n}(z_j - z_k^*)\right| = \left|\prod_{k=1}^{n}\prod_{j=1}^{d}(z_k^* - z_j)\right|$$

we have

$$\left|\prod_{j=1}^{d} T_n^*(z_j)\right| = \left|\prod_{k=1}^{n}(h(z_k^*) - w_0)\right| = |q_n(w_0)|,$$

where

$$q_n(w) = \prod_{k=1}^{n}(w - h(z_k^*)).$$

Since $h(z_k^*) \in E$, we can deduce that

$$m_n \le \max_{w \in E} |q_n(w)| \le \left(\max_{z \in E^*} |T_n^*(z)|\right)^d = (m_n^*)^d,$$

showing that

$$\left(m_n^{1/n}\right)^{1/d} \le (m_n^*)^{1/n},$$

and therefore that $\tau(E)^{1/d} \le \tau(E^*)$. □

Proposition 2.8. *Define E_ϵ, the ϵ-thickening of E, to be the set of all $z \in \mathbb{C}$ distant at most ϵ from some element of E. Then the limit $\lim_{\epsilon \searrow 0} \tau(E_\epsilon)$ exists and equals $\tau(E)$.*

Proof. We know that $\tau(E) \le \tau(E_\epsilon)$. Given $\delta > 0$, we will show that $\tau(E_\epsilon) \le \tau(E) + 2\delta$ for all small enough $\epsilon > 0$. Put $\nu = n(n-1)/2$ and suppose that

$$G_n(E_\epsilon) = \left|\prod_{\substack{j,k=1 \\ j<k}}^{n}(z_j^* - z_k^*)\right|^{1/\nu},$$

where $z_j^* = z_j + \mu_j\epsilon \in E_\epsilon$, with $z_j \in E$ and $|\mu_j| \le 1$. Then

$$\begin{aligned} G_n(E_\epsilon)^\nu &= \left| \prod_{\substack{j,k=1 \\ j<k}}^{n} (z_j - z_k + (\mu_j - \mu_k)\epsilon) \right| \\ &\le \prod_{\substack{j,k=1 \\ j<k}}^{n} (|z_j - z_k| + 2\epsilon) \\ &= \prod_{\substack{j,k=1 \\ j<k}}^{n} |z_j - z_k| + \prod_{\substack{j,k=1 \\ j<k}}^{n} (|z_j - z_k| + 2\epsilon) - \prod_{\substack{j,k=1 \\ j<k}}^{n} |z_j - z_k| \\ &\le \prod_{\substack{j,k=1 \\ j<k}}^{n} |z_j - z_k| + (D + 2\epsilon)^\nu - D^\nu, \end{aligned}$$

where we recall that D is the diameter of E. For the last line, we used Lemma 2.10. So since the sequence $G_n(E_\epsilon)$ is monotonically decreasing, we have

$$\tau(E_\epsilon) \le (G_n^\nu + (D + 2\epsilon)^\nu - D^\nu)^{1/n}.$$

Upon choosing n sufficiently large so that $G_n \le \tau(E) + \delta$, we then have

$$\begin{aligned} \tau(E_\epsilon) &\le ((\tau(E) + \delta)^\nu + (D + 2\epsilon)^\nu - D^\nu)^{1/\nu} \\ &\le (\tau(E) + \delta) \left(1 + \frac{(D + 2\epsilon)^\nu - D^\nu}{\delta^\nu} \right)^{1/\nu}, \end{aligned}$$

using $\tau(E) + \delta \ge \delta$. Then we can choose $\epsilon > 0$ small enough so that

$$\left(1 + \frac{(D + 2\epsilon)^\nu - D^\nu}{\delta^\nu} \right)^{1/\nu} \le 1 + \frac{\delta}{\tau(E) + \delta},$$

giving $\tau(E_\epsilon) \le \tau(E) + 2\delta$, as we wanted. □

Proposition 2.9. *The boundary ∂E of E has the same transfinite diameter as E.*

Proof. As earlier, let $V_n(E)$ be the maximum of $|V(z_1, \dots, z_n)|$ as $z_1, \dots z_n$ range over E. For any particular z_j, the maximum occurs for z_j on the boundary of E, by the maximum principle. Hence $V_n(E) = V_n(\partial E)$, and so $\tau(E) = \tau(\partial E)$. □

Lemma 2.10. *Let δ and $a_1, \dots, a_n$ be positive numbers with $A := \max_{j=1}^n a_j$. Then*

$$\prod_{j=1}^{n}(a_j+\delta) - \prod_{j=1}^{n} a_j \le (A+\delta)^n - A^n.$$

Proof. We have

$$\begin{aligned}\prod_{j=1}^{n}(a_j+\delta) &= \sum_{B\subset\{1,2,\dots,n\}} \delta^{n-|B|}\prod_{b\in B} a_b \\ &= \sum_{B\subsetneq\{1,2,\dots,n\}} \delta^{n-|B|}\prod_{b\in B} a_b + a_1a_2\cdots a_n \\ &\le \sum_{B\subsetneq\{1,2,\dots,n\}} \delta^{n-|B|}\prod_{b\in B} A + a_1a_2\cdots a_n \\ &= \sum_{B\subsetneq\{1,2,\dots,n\}} \delta^{n-|B|}\cdot A^{|B|} + a_1a_2\cdots a_n.\end{aligned}$$

Now because for every $k \in \{1,2,\dots,n\}$ there are $\binom{n}{k}$ subsets of $\{1,2,\dots,n\}$ with k elements we have that

$$\prod_{j=1}^{n}(a_j+\delta) \le \binom{n}{0}\delta^n + \binom{n}{1}\delta^{n+1}A^1 + \cdots + \binom{n}{n-1}\delta^1A^{n-1} + a_1a_2\cdots a_n.$$

Hence

$$\begin{aligned}\prod_{j=1}^{n}(a_j+\delta) - \prod_{j=1}^{n} a_j &\le \binom{n}{0}\delta^n + \cdots + \binom{n}{n-1}\delta^1A^{n-1} + a_1a_2\cdots a_n - \prod_{j=1}^{n} a_j \\ &= \left(\binom{n}{0}\delta^n + \cdots + \binom{n}{n-1}\delta^1A^{n-1} + A^n\right) - A^n \\ &= (A+\delta)^n - A^n.\end{aligned}$$

□

2.2 Example: the unit circle

Now, for a practical example we will compute the transfinite diameter of the unit circle. We will try two different methods, first we will compute it using Definition 2.1 and then using Proposition 2.6.

Proposition 2.11. *Let E be the unit circle in $\mathbb{C}$. Then $\tau(E) = 1$.*

Proof. We use Euler's formula

$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$

The point $e^{i\theta}$ is the point on the unit circle that is rotated θ radians from standard position. The distance between two points z_k and z_j on the unit circle with $\arg(z_k) = \theta_k$ and $\arg(z_j) = \theta_j$ is given by:

$$|z_k - z_j| = |e^{i\theta_k} - e^{i\theta_j}|.$$

We claim that the products of the distances between n points is maximized when the points are equally spaced around the unit circle, in the shape of a regular n-gon.

Consider n points on the unit circle represented as $z_k = e^{i\theta_k}$ for $k = 0, 1, \ldots, n-1$. We want to maximize the product of distances between all pairs of these points

$$p_n(\theta_0, \theta_1, \ldots, \theta_{n-1}) := \prod_{0 \le k < j \le n-1} \left| e^{i\theta_k} - e^{i\theta_j} \right|$$

The distance is

$$\begin{aligned} \left| e^{i\theta_k} - e^{i\theta_j} \right| &= |(\cos(\theta_k) + i\sin(\theta_k)) - (\cos(\theta_j) + i\sin(\theta_j))| \\ &= |\cos(\theta_k) - \cos(\theta_j) + i(\sin(\theta_k) - \sin(\theta_j))| \end{aligned}$$

The modulus of a complex number $a + ib$ is equal to $\sqrt{a^2 + b^2}$, thus

$$\begin{aligned} |e^{i\theta_k} - e^{i\theta_j}| &= \sqrt{(\cos(\theta_k) - \cos(\theta_j))^2 + (\sin(\theta_k) - \sin(\theta_j))^2} \\ &= \big(\cos^2(\theta_k) - 2\cos(\theta_k)\cos(\theta_j) + \cos^2(\theta_j) \\ &\qquad + \sin^2(\theta_k) - 2\sin(\theta_k)\sin(\theta_j) + \sin^2(\theta_j) \big)^{1/2} \\ &= \sqrt{1 - 2\cos(\theta_k)\cos(\theta_j) + 1 - 2\sin(\theta_k)\sin(\theta_j)} \\ &= \sqrt{2 - 2(\cos(\theta_k)\cos(\theta_j) + \sin(\theta_k)\sin(\theta_j))}. \end{aligned}$$

Since $\cos(\theta_k - \theta_j) = \cos(\theta_k)\cos(\theta_j) + \sin(\theta_k)\sin(\theta_j)$, we get

$$|e^{i\theta_k} - e^{i\theta_j}| = \sqrt{2 - 2\cos(\theta_k - \theta_j)}.$$

Using the double-angle identity for cosine $1 - \cos(x) = 2\sin^2\left(\frac{x}{2}\right)$ for $x = \theta_k - \theta_j$, we get

$$2 - 2\cos(\theta_k - \theta_j) = 2 \cdot 2\sin^2\left(\frac{\theta_k - \theta_j}{2}\right) = 4\sin^2\left(\frac{\theta_k - \theta_j}{2}\right).$$

Therefore,

$$\left|e^{i\theta_k} - e^{i\theta_j}\right| = 2\sin\left(\frac{\theta_k - \theta_j}{2}\right).$$

Thus, the product to maximize becomes

$$p_n(\theta_0, \theta_1, \ldots, \theta_{n-1}) = \prod_{0 \le k < j \le n-1} 2\sin\left(\frac{\theta_k - \theta_j}{2}\right).$$

Since $2^{\binom{n}{2}}$ is a constant factor, we can focus on maximizing

$$q(\theta_0, \theta_1, \ldots, \theta_{n-1}) = \prod_{0 \le k < j \le n-1} \sin\left(\frac{\theta_k - \theta_j}{2}\right).$$

We need to find the critical points of q. To do this, we take the natural logarithm to convert the product into a sum

$$\log q(\theta_0, \theta_1, \ldots, \theta_{n-1}) = \sum_{0 \le k < j \le n-1} \log\left(\sin\left(\frac{\theta_k - \theta_j}{2}\right)\right).$$

Let $f(\theta_0, \theta_1, \ldots, \theta_{n-1}) = \log q(\theta_0, \theta_1, \ldots, \theta_{n-1})$. Then we need to find the critical points of f. The partial derivative of f with respect to θ_m is

$$\frac{\partial f}{\partial \theta_m} = \sum_{k \neq m} \frac{\partial}{\partial \theta_m} \log\left(\sin\left(\frac{\theta_m - \theta_k}{2}\right)\right).$$

Using the chain rule, we get

$$\frac{\partial}{\partial \theta_m} \log\left(\sin\left(\frac{\theta_m - \theta_k}{2}\right)\right) = \frac{1}{\sin\left(\frac{\theta_m - \theta_k}{2}\right)} \cdot \frac{\cos\left(\frac{\theta_m - \theta_k}{2}\right)}{2} = \frac{1}{2}\cot\left(\frac{\theta_m - \theta_k}{2}\right).$$

So

$$\frac{\partial f}{\partial \theta_m} = \frac{1}{2}\sum_{k \neq m} \cot\left(\frac{\theta_m - \theta_k}{2}\right).$$

To solve the equation

$$\sum_{k \neq m} \cot\left(\frac{\theta_m - \theta_k}{2}\right) = 0 \tag{2.4}$$

we need to determine the values of θ_k such that this condition holds for all m.

Consider n points equally spaced on the unit circle. The angles θ_k for these points can be written as

$$\theta_k = \frac{2\pi k}{n} \quad \text{for} \quad k = 0, 1, 2, \ldots, n-1.$$

These points are separated by an equal angle of $\frac{2\pi}{n}$. Substitute $\theta_k = \frac{2\pi k}{n}$ into the sum (2.4)

$$\sum_{k \neq m} \cot\left(\frac{\frac{2\pi m}{n} - \frac{2\pi k}{n}}{2}\right) = \sum_{k \neq m} \cot\left(\frac{\pi(m-k)}{n}\right).$$

We need to show that:

$$\sum_{k \neq m} \cot\left(\frac{\pi(m-k)}{n}\right) = 0.$$

We observe that set of angles $\theta_k = \frac{2\pi k}{n}$ is symmetric. Also note that the cotangent function, $\cot(x)$, is odd and periodic with period π

$$\cot(x+\pi) = \cot(x) \quad \text{and} \quad \cot(-x) = -\cot(x).$$

Thus the sum of cotangents over equally spaced points will involve terms that are symmetric about m. For each $k \neq m$, there exists a $k' \neq m$ such that $\cot\left(\frac{\pi(m-k)}{n}\right)$ is paired with $\cot\left(\frac{\pi(m-k')}{n}\right)$. These pairs of cotangents have equal absolute values but opposite signs. Therefore the sum adds up to zero and this proves our claim.

We consider these points that maximize the function p_n as if they were in the complex plane. Let $z_1 = 1 = e^{i \cdot 0}$. According to our previous calculation, the next point in our collection will be rotated $\frac{2\pi}{n}$ radians from 1. Thus the second point is $e^{i\frac{2\pi}{n}}$. Continuing in this manner, the other points that are equally spaced around the unit circle are the other complex nth roots of unity, and every nth root of unity will be in our collection. So, every point in our collection is a root of the equation

$$x^n - 1 = 0.$$

Thus,

$$x^n - 1 = (x-1)(x - e^{i\frac{2\pi}{n}})(x - e^{i\frac{4\pi}{n}})\cdots(x - e^{i\frac{(2n-2)\pi}{n}})$$

However, we also have

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1).$$

and dividing out the $(x-1)$ term yields

$$(x - e^{i\frac{2\pi}{n}})(x - e^{i\frac{4\pi}{n}})\cdots(x - e^{i\frac{2n-2\pi}{n}}) = (x^{n-1} + x^{n-2} + \cdots + x + 1)$$

Substituting 1 for x in each polynomial gives:

$$(1 + 1 + \cdots + 1) = n = (1 - e^{i\frac{2\pi}{n}})(1 - e^{i\frac{4\pi}{n}})\cdots(1 - e^{i\frac{(2n-2)\pi}{n}})$$

This implies that

$$\begin{aligned} n &= |(1 - e^{i\frac{2\pi}{n}})(1 - e^{i\frac{4\pi}{n}})\cdots(1 - e^{i\frac{(2n-2)\pi}{n}}| \\ &= |(1 - e^{i\frac{2\pi}{n}})||(1 - e^{i\frac{4\pi}{n}})|\cdots|(1 - e^{i\frac{(2n-2)\pi}{n}}| \end{aligned}$$

We denote this product of the distances from $z_1 = 1$ to each other point by $C(z_1)$. By the symmetry of the points, the $C(z_k) = C(z_j)$ for $1 \leq k, j \leq n$. Thus we have

$$\prod_{j=0}^{n-1} C(e^{i\frac{2j\pi}{n}}) = n^n.$$

Given any two points z_j and z_k, the distance between them appears twice in the preceding product, once in $C(z_j)$ and once in $C(z_k)$. Thus we have

$$\begin{aligned} p_n^2 &= n^n \\ p_n &= n^{\frac{n}{2}} \\ p_n^{\frac{2}{n(n-1)}} &= G_n = n^{\frac{1}{n-1}} \end{aligned}$$

and so because $e^{\frac{\log(n)}{n-1}} \to 1$ as $n \to \infty$, we obtain

$$\tau(E) = \lim_{n\to\infty} G_n = 1.$$

□

Now we give an alternative proof using Chebyshev polynomials. Before we begin we recall that for E a compact set in the complex plane $\mathbb{C}$, the uniform (sup) norm on E is defined as

$$||f||_E = \sup_{z \in E} |f(z)|.$$

Note that for $T_n(z)$ the nth Chebyshev polynomial we have $||T_n||_E = m_n$ (see Definition 2.2).

Proof. (second proof of Proposition 2.11) Let $E = \overline{D(0,1)} = \{z \in \mathbb{C} : |z| \leq 1\}$ be the closed disk with centre at 0 and radius 1. We consider some monic polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0 \in \mathbb{C}[z]$. By Proposition 2.9, the transfinite diameter of the closed disk is the same as the transfinite diameter of the unit circle. Observe that

$$\begin{aligned} ||p||_E^2 \geq ||p||_{L^2}^2 &= \frac{1}{2\pi} \int_0^{2\pi} |p(e^{it})|^2 dt \\ &= 1 + |a_{n-1}|^2 + \cdots + |a_0|^2 \geq 1, \end{aligned}$$

where the second equality comes from Parseval's identity and the fact that the Fourier coefficients of $p(e^{it})$ are the coefficients of the polynomial p. So $||p||_E \geq 1$ for any monic polynomial. Now observe that $||z^n||_E = 1$ Therefore for the Chebyshev polynomial must satisfy $1 \leq ||T_n(z)||_K = m_n \leq 1$. So $m_n = 1$ and $\tau_n = 1$ and thus $\tau = 1$. □

2.3 Example: the unit segment

Consider the line segment $E = \{x + 0 \cdot i \mid x \in [0,1]\}$. For simplicity we consider it as a subset of the real numbers, $E = [0,1] \in \mathbb{R}$. Unlike in the case of the unit circle the obvious choice of equally distanced points does not give the actual values for $G_n(E)$, i.e., $\prod_{1 \leq j < k \leq n} |z_j - z_k|$ is not maximized for $z_k = \frac{k-1}{n}$, but they give us the bound $e^{-3/2} \approx 0.22313 \leq \tau(E)$. In fact as we will prove later $\tau(E) = \frac{1}{4}$. First we will analyse what happens when we consider the equally distanced points and then we will compute the transfinite diameter using the equivalent definition by Proposition 2.6.

Just like in the unit circle we consider the points that are equally spaced along the segment. Set $p(z_1, \ldots, z_n) := \prod_{1 \leq j < k \leq n} |z_j - z_k|$ and consider the $n+1$ points $\{0, \frac{1}{n}, \frac{2}{n}, \ldots \frac{n}{n}\}$. We will consider the limit of this value as n goes to infinity to give a lower bound for $\tau([0,1])$, since the points are not actually the optimal points. The product $p\left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right)$ contains $n(n+1)/2$ terms and each of these terms will be a fraction with denominator n. The numerators of these fractions will contain n values of 1, $n-1$ values of 2, etc., up to 1 value of n. Thus, we have

$$p\left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right) = \frac{n!(n-1)!(n-2)!\cdots 2!1!}{n^{\frac{n(n+1)}{2}}}.$$

From [3] we have that

$$\lim_{n\to\infty} p\left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right)^{\frac{2}{(n+1)n}} = e^{-3/2} \approx 0.22313.$$

Therefore $p\left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right)^{\frac{2}{(n+1)n}} \leq G_{n+1}(E)$, we take the limit as n goes to infinity of both sides to get

$$e^{-3/2} \approx 0.22313 \leq \tau(E).$$

Before we continue note that we can assume that the Chebychev polynomials of the closed interval $[-2,2]$ have real coefficients. To see this let $p(x)$ be any monic polynomial with complex coefficients and let $x \in [-2,2]$. Then since x is real we have $|p(x)| = |\overline{p(x)}| = |\overline{p}(x)|$. Now consider the polynomial $r(x) := \frac{p(x)+\overline{p(x)}}{2}$ and note that it is monic and has real coefficients. Then we have

$$|r(x)| = \frac{|p(x)+\overline{p}(x)|}{2} \leq \frac{|p(x)|+|\overline{p}(x)|}{2} = |p(x)|.$$

Proposition 2.12. *Let E be the line segment $E = \{x+0\cdot i \mid x \in [-2,2]\} \in \mathbb{C}$. Then $\tau(E) = 1$.*

Proof. As above we consider E as a subset of the real numbers, $E = [-2,2] \in \mathbb{R}$. Consider some monic polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0 \in \mathbb{R}[z]$.

We parametrise by $z = 2\cos t$ and set

$$q(t) := p(2\cos t) = (2\cos t)^n + a_{n-1}(\cos t)^{n-1} + \cdots + a_0,\ t \in [0, 2\pi].$$

We observe that

$$||p||_E = \max_{t\in[0,2\pi]} |p(2\cos t)| = \max_{t\in[0,2\pi]} |q(t)|$$

and also using trigonometric identities for trigonometric powers we observe that $q(t) = 2\cos(nt) +$ "some polynomial in $cos(jt)$ with $j \leq n-1$". Now similarly to the case of the unit circle we use Parseval's identity and obtain

$$||q(t)||_K^2 \geq ||q(t)||_{L^2}^2 = \frac{1}{2\pi}\int_0^{2\pi} |q(t)|^2 dt$$
$$= 4 + \text{"sum of positive real numbers"} \geq 4.$$

For a complex number $z = e^{it}$ now consider the Dickson polynomials [15]

$$D_n(z + z^{-1}) = z^n + z^{-n} = e^{int} + e^{-int} = 2\cos(nt).$$

We have that $||D_n||_E = 2$. Thus $m_n = 2$ and $\tau_n = 2^{1/n}$, so $\tau_E = 1$. □

By the scaling property of the transfinite diameter we obtain the following.

Corollary 2.13. *Let E be the line segment $E = \{x + 0 \cdot i \mid x \in [0,1]\} \in \mathbb{C}$. Then $\tau(E) = 1/4$.*

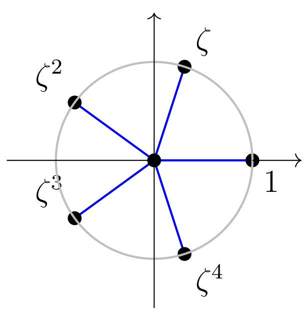

Figure 3: The regular 5-hedgehog $\mathcal{H}_5 := \mathcal{H}(1, \zeta, \ldots, \zeta^4)$, $\zeta = e^{2\pi i/5}$.

Corollary 2.14. *Let $\mathcal{H}_n := \mathcal{H}(1, \zeta_n, \ldots, \zeta_n^{n-1})$, $\zeta_n = e^{2\pi i/n}$ be the regular n-hedgehog. Then $\tau(\mathcal{H}_n) = 4^{-1/n}$.*

Proof. We have that $\mathcal{H}_n = f^{-1}[0,1]$ where $f(x) = x^n$. Hence by Proposition 2.7 $\tau(\mathcal{H}_n) = 4^{-1/n}$. □

3 Bounds on determinants

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

$$A_N = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_0 & X_1 & \cdots & X_m \\ X_0^2 & X_1^2 & \cdots & X_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ X_0^{N-1} & X_1^{N-1} & \cdots & X_m^{N-1} \end{pmatrix} \in \mathrm{Mat}_{N,m+1}(\mathbb{Z}[X_0, \ldots, X_m]). \quad (3.1)$$

Note that for $N = m+1$ we recover a Vandermonde matrix.

The goal of this section is to bound the quantities

$$X = \frac{1}{2} \log \det A_N(\overline{z_0}, \ldots, \overline{z_m})^\top A_N(z_0, \ldots, z_m)),$$

where $z_0, \ldots, z_m \in \mathbb{C}$. The obtained bounds (Proposition 3.4) are key ingredients for bounding transfinite diameters of hedgehogs in the next section.

3.1 Statement of result

Definition 3.1. For a real number t, we define $\log^+ t = \log \max\{1, t\}$.

Definition 3.2. Let $m \in \mathbb{Z}$, the Barnes G-function is defined as

$$G(m+2) = \begin{cases} 0 & m \leq -2 \\ 1 & m = -1, 0 \\ 1!2! \cdots m! & m \geq 1 \end{cases}$$

Definition 3.3. Let N and m be integers with $N \geq m+1$ and $m \geq -1$. We define

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

Proposition 3.4. *Let $N \geq 2$ and $m \geq 0$ be integers with $N \geq m+1$ and let $z_0, \ldots, z_m \in \mathbb{C}$ be pairwise distinct. Set*

$$X = \frac{1}{2} \log \left(\det A_N(\overline{z_0}, \ldots, \overline{z_m})^\top A_N(z_0, \ldots, z_m) \right).$$

(i) We have

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

(ii) We have

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

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

$$X \leq (m+1)\log N + \frac{m(m+1)}{2}\log\max\{|z_0|,\ldots,|z_m|\}, \tag{3.2}$$

the right-hand side is taken to be $\log N$ *in the case* $m=0$*.*

Corollary 3.5. *With the same notations as above, we have*

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

3.2 Preliminary results

Proposition 3.6. *Let* $z_0,\ldots,z_m \in \mathbb{C}$ *be pairwise distinct, then*

$$\det\left(A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m)\right) \neq 0.$$

Proof. The vectors $(1, z_i, z_i^2, \ldots, z_i^{N-1})$ for $0 \leq i \leq m$ are the columns of the matrix $A_N(z_0,\ldots,z_m)$ thus the matrix $A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m)$ is a Gram matrix. The determinant of the Gram matrix is non-zero if and only if the vectors $(1, z_i, z_i^2, \ldots, z_i^{N-1})$, $0 \leq i \leq m$ are linearly independent. To show that these vectors are linearly independent we first choose complex numbers

z_i for $m+1 \leq N-1$ such that the numbers $z_0, z_1, \ldots, z_{N-1}$ are pairwise distinct. Now consider the matrix

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_0 & z_1 & \cdots & z_{N-1} \\ z_0^2 & z_1^2 & \cdots & z_{N-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_0^{N-1} & z_1^{N-1} & \cdots & z_{N-1}^{N-1} \end{pmatrix}.$$

The determinant of this matrix is $\prod_{0\leq i<j\leq N-1}(z_j - z_i)$ by Proposition A.3 and it is non-zero since the z_i are distinct. Thus the vectors $(1, z_i, z_i^2, \ldots, z_i^{N-1})$ for $0 \leq i \leq N-1$ are linearly independent and therefore the vectors $(1, z_i, z_i^2, \ldots, z_i^{N-1})$ for $0 \leq i \leq m$ are linearly independent. □

Before we prove the Proposition 3.4 we need a few definitions and a few lemmas.

Definition 3.7. Let $m \in \mathbb{N}$ and let $I = (\alpha_0, \ldots, \alpha_m)$ be an $(m+1)$-tuple of non-negative and strictly increasing integers. We define

$$A_I = \begin{pmatrix} X_0^{\alpha_0} & X_1^{\alpha_0} & \cdots & X_m^{\alpha_0} \\ X_0^{\alpha_1} & X_1^{\alpha_1} & \cdots & X_m^{\alpha_1} \\ \vdots & \vdots & \vdots & \vdots \\ X_0^{\alpha_m} & X_1^{\alpha_m} & \cdots & X_m^{\alpha_m} \end{pmatrix}. \tag{3.3}$$

Definition 3.8. Given $j \in \mathbb{N}_0$ and $k \in \mathbb{Z}$ the complete homogeneous symmetric polynomial of degree k in j variables is

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

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

Lemma 3.9 (Jacobi-Trudi identity)**.** *Given $I = (\alpha_0, \ldots, \alpha_m)$ be as above , we define*

$$S_I = det \begin{pmatrix} h_{\alpha_0}(X_0) & h_{\alpha_0-1}(X_0, X_1) & \cdots & h_{\alpha_0-m}(X_0, \ldots, X_m) \\ h_{\alpha_1}(X_0) & h_{\alpha_1-1}(X_0, X_1) & \cdots & h_{\alpha_1-m}(X_0, \ldots, X_m) \\ \vdots & \vdots & \ddots & \vdots \\ h_{\alpha_m}(X_0) & h_{\alpha_m-1}(X_0, X_1) & \ldots & h_{\alpha_m-m}(X_0, \ldots, X_m) \end{pmatrix}, \tag{3.4}$$

Then

$$\det A_I = S_I \prod_{0\le i,j\le m} (X_j - X_i). \tag{3.5}$$

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

$$det \begin{pmatrix} h_{\alpha_0}(X_0) & \cdots & h_{\alpha_0-k}(X_0,\dots,X_{k-1},X_k) & \cdots & h_{\alpha_0-k}(X_0,\dots,X_{k-1},X_m) \\ h_{\alpha_1}(X_0) & \cdots & h_{\alpha_1-k}(X_0,\dots,X_{k-1},X_k) & \cdots & h_{\alpha_1-k}(X_0,\dots,X_{k-1},X_m) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{\alpha_m}(X_0) & \dots & h_{\alpha_m-k}(X_0,\dots,X_{k-1},X_k) & \cdots & h_{\alpha_m-k}(X_0,\dots,X_{k-1},X_m) \end{pmatrix}. \tag{3.6}$$

The lemma then follows by taking $k = m$. We proceed by induction.

The claim holds true for $k = 0$ as

$$A_I = \begin{pmatrix} X_0^{\alpha_0} & X_1^{\alpha_0} & \cdots & X_m^{\alpha_0} \\ X_0^{\alpha_1} & X_1^{\alpha_1} & \cdots & X_m^{\alpha_1} \\ \vdots & \vdots & \ddots & \vdots \\ X_0^{\alpha_m} & X_1^{\alpha_m} & \cdots & X_m^{\alpha_m} \end{pmatrix}.$$

We assume that the claim holds for $k \in \{0,\dots,m-1\}$. The matrix in (3.6) 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

$$\det A_I = \det(c_0,\dots,c_k,c_{k+1},c_{k+1}-c_k,c_{k+2}-c_k,\dots,c_m-c_k)\prod_{i=0}^{k-1}\prod_{j=i+1}^{m}(X_j-X_i). \tag{3.7}$$

Indeed for $j \in \{1,\dots m-k\}$

$$c_{k+j} - c_k = \begin{pmatrix} h_{\alpha_0-k}(X_0,\dots,X_{k-1},X_{k+j}) - h_{\alpha_0-k}(X_0,\dots,X_k) \\ h_{\alpha_1-k}(X_0,\dots,X_{k-1},X_{k+j}) - h_{\alpha_1-k}(X_0,\dots,X_k) \\ \vdots \\ h_{\alpha_m-k}(X_0,\dots,X_{k-1},X_{k+j}) - h_{\alpha_m-k}(X_0,\dots,X_k) \end{pmatrix}$$

with

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

So we can factor out $X_{k+j} - X_k$ from each respective column. We insert this into (3.7) and find that $\det A_I$ equals

$$\det\begin{pmatrix} c_0 & \cdots & c_k & \begin{matrix} h_{\alpha_0-(k+1)}(X_0,\ldots,X_k,X_{k+1}) & \ldots & h_{\alpha_0-(k+1)}(X_0,\ldots,X_k,X_m)\\ h_{\alpha_1-(k+1)}(X_0,\ldots,X_k,X_{k+1}) & \ldots & h_{\alpha_1-(k+1)}(X_0,\ldots,X_k,X_m)\\ \vdots & \ddots & \vdots\\ h_{\alpha_m-(k+1)}(X_0,\ldots,X_k,X_{k+1}) & \ldots & h_{\alpha_m-(k+1)}(X_0,\ldots,X_k,X_m)\end{matrix}\end{pmatrix}$$

$$\cdot\left(\prod_{i=0}^{k-1}\prod_{j=i+1}^{m}(X_j-X_i)\right)\prod_{j=1}^{m-k}(X_{k+j}-X_k).$$

Observing that

$$\left(\prod_{i=0}^{k-1}\prod_{j=i+1}^{m}(X_j-X_i)\right)\prod_{j=1}^{m-k}(X_{k+j}-X_k) = \prod_{i=0}^{k}\prod_{j=i+1}^{m}(X_j-X_i)$$

finishes the proof. □

Lemma 3.10. *Let I and S_I be as above. The coefficients of S_I are non-negative integers.*

Proof. From the Lemma 3.9 we have that

$$S_I = \frac{\det A_I}{\prod_{0\le i,j\le m}(X_j - X_i)} = \frac{\det A_I}{\det V(X_0,\ldots,X_m)}$$

where $V(X_0,\ldots,X_m)$ is a Vandermonde matrix. Thus by Proposition A.9 we have that

$$S_I = s_\lambda(x_1,x_2,\ldots,x_n) = \sum_T x^{w(T)},$$

where the sum runs over all tableaux of shape λ. The coefficients of $\sum_T x^{w(T)}$ are non negative, finishing the proof. □

Lemma 3.11. *Let $D \in Mat_{m+1}(\mathbb{C}[T])$ and $t \in \mathbb{C}$ such that the rank of $D(t) \in Mat_{m+1}(\mathbb{C})$ 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 $U, V \in \mathrm{GL}_{m+1}(\mathbb{C}[T])$ such that UDV is diagonal with diagonal entries $f_0, \ldots, f_m \in \mathbb{C}[T]$. 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 $f_0, \ldots, f_m$ at t. The lemma follows, as by hypothesis at most r among $f_0(t), \ldots, f_m(t)$ are non-zero. □

Lemma 3.12. *Suppose $N \geq 2$ and $m \geq 0$ are integers with $N \geq m+1$. Then*

$$\det A_N(X_0, \ldots, X_m)^\top A_N(Y_0, \ldots, Y_m) = B \prod_{0 \leq i < j \leq m} (X_j - X_i)(Y_j - Y_i) \quad (3.8)$$

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

$$B(\underbrace{1, \ldots, 1}_{2m+2 \text{ times}}) = \gamma_{N,m} N^{m+1}. \quad (3.9)$$

Moreover, $\max_{0 \leq j \leq m}\{\deg_{X_j} B, \deg_{Y_j} B\} \leq N - (m+1)$.
Note that for $N = m+1$ we are in the Vandermonde case and (3.9) implies

$$\gamma_{N,N-1} = N^{-N}. \quad (3.10)$$

Proof. By the Cauchy-Binet Formula (Theorem A.2), the left-hand side of (3.8) equal

$$\sum_I \det A_I(X_0, \ldots, X_m)^\top A_I(Y_0, \ldots, Y_m)$$

where here and below the sum ranges over all $(m+1)$-tuples $I = (\alpha_0, \ldots \alpha_m)$ of integers satisfying $0 \leq \alpha_0 < \cdots < \alpha_m \leq N-1$. Now Lemma 3.9 implies (3.8) with

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

and with S_I as in (3.4). Thus $B \in \mathbb{Z}[X_0, \ldots, X_m, Y_0, \ldots, Y_m]$. Moreover, each S_I has non-negative coefficients by Lemma 3.10 and thus the same holds for B.

The degree of $\det A_N(X_0, \dots, X_m)^\top A_N(Y_0, \dots, Y_m)$ with respect to X_j is at most $N-1$. The degree of the Vandermonde determinant $\prod_{0 \le i < j \le m}(X_j - X_i)$ with respect to X_j is m. So (3.8) implies $\deg_{X_j} B \le N - (m+1)$ and the same bound holds for $\deg_{Y_i} B$.

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

$$S_I = \begin{pmatrix} h_{\alpha_0}(1) & h_{\alpha_0 - 1}(1,1) & \cdots & h_{a_0 - m}(1, \dots, 1) \\ h_{\alpha_1}(1) & h_{\alpha_1 - 1}(1,1) & \cdots & h_{a_1 - m}(1, \dots, 1) \\ \vdots & \vdots & \ddots & \vdots \\ h_{\alpha_m}(1) & h_{\alpha_m - 1}(1,1) & \cdots & h_{a_m - m}(1, \dots, 1) \end{pmatrix}.$$

Also observe that $h_{a_i - j}(\underbrace{1, \dots, 1}_{j+1 \text{ times}})$ is equal to the number of ways one can partition $\alpha_i - j$ in $j+1$ parts and

$$\frac{(\alpha_j - j + j + 1 - 1)!}{(j + 1 - 1)!(\alpha_j - j)!} = \frac{\alpha_i!}{j!(\alpha_i - j)!} = \binom{\alpha_i}{j},$$

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

$$B(1, \dots, 1) = \sum_I b_I^2$$

where

$$b_{(\alpha_0, \dots, \alpha_m)} = \det \begin{pmatrix} \binom{\alpha_0}{0} & \binom{\alpha_0}{1} & \cdots & \binom{\alpha_0}{m} \\ \binom{\alpha_1}{0} & \binom{\alpha_1}{1} & \cdots & \binom{\alpha_1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{\alpha_m}{0} & \binom{\alpha_m}{1} & \cdots & \binom{\alpha_m}{m} \end{pmatrix}.$$

Observe that

$$b_{(\alpha_0, \dots, \alpha_m)} = \frac{1}{1!2! \cdots m!} \det \begin{pmatrix} 1 & \alpha_0 & \alpha_0(\alpha_0 - 1) & \cdots & \alpha_0(\alpha_0 - 1) \cdots (\alpha_0 - m + 1) \\ 1 & \alpha_1 & \alpha_1(\alpha_1 - 1) & \cdots & \alpha_1(\alpha_1 - 1) \cdots (\alpha_1 - m + 1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_m & \alpha_m(\alpha_m - 1) & \cdots & \alpha_m(\alpha_m - 1) \cdots (\alpha_m - m + 1) \end{pmatrix}$$

An entry in the $(j+1)$-st column of the above matrix is of the form $\alpha_i^j +$ (polynomial in α_i of degree $< j$). So by Proposition A.4

$$b_{(\alpha_0,\ldots,\alpha_m)} = \frac{1}{G(m+2)} \det \begin{pmatrix} \alpha_0^0 & \alpha_0^1 & \cdots & \alpha_0^m \\ \alpha_1^0 & \alpha_1^1 & \cdots & \alpha_1^m \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_m^0 & \alpha_m^1 & \cdots & \alpha_m^m \end{pmatrix},$$

where G is the Barnes G-function from Definition 3.2 and where we use the convention $0^0 = 1$. Thus by the Cauchy-Binet Formula (Theorem A.2)

$$B(1,\ldots,1) = \sum_I b_I^2 = \frac{\det C^\top C}{G(m+2)^2} \tag{3.11}$$

with

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & N-1 & \cdots & (N-1)^m \end{pmatrix},$$

where

$$C^\top C = \left(\sum_{k=0}^{N-1} k^{i+j} \right)_{0 \le i,j \le m} = \begin{pmatrix} \sum_{k=0}^{N-1} k^{0+0} & \sum_{k=0}^{N-1} k^{0+1} & \cdots & \sum_{k=0}^{N-1} k^{0+m} \\ \sum_{k=0}^{N-1} k^{1+0} & \sum_{k=0}^{N-1} k^{1+1} & \cdots & \sum_{k=0}^{N-1} k^{1+m} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} k^{m+0} & \sum_{k=0}^{N-1} k^{m+1} & \cdots & \sum_{k=0}^{N-1} k^{m+m} \end{pmatrix}.$$

Note that the top-left entry is N.

We now prove that

$$\det C^\top C = \gamma_{N,m} N^{m+1} G(m+2)^2. \tag{3.12}$$

This equation together with (3.11) implies (3.9) and completes the proof.

Let $i \ge 0$ be an integer and let $B_i = T^i + \cdots \in \mathbb{Q}[T]$ denote the i-th Bernoulli polynomial with constant term $B_i(0)$. Recall Faulhaber's Formula (Proposition A.13):

$$\sum_{k=0}^{N-1} k^i = s_i(N) \text{ where } s_i = \frac{(B_{i+1}(T) - B_{i+1}(0))}{i+1} \in \mathbb{Q}[T].$$

We define

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

with $D \in \mathrm{Mat}_{m+1}(\mathbb{Q}[T])$ and find $C^\top C = D(N)$. The determinant is $\det D = \sum_\sigma \mathrm{sign}(\sigma) s_{0+\sigma(0)} s_{1+\sigma(1)} \cdots s_{m+\sigma(m)}$ where σ ranges over all permutations of $\{0, 1, \ldots, m\}$. As $\deg s_{i+\sigma(i)} = i + \sigma(i) + 1$ we find that $\deg \det D \le (m+1)^2$. Because for any of the σ's

$$\begin{aligned} &\deg(s_{0+\sigma(0)} s_{1+\sigma(1)} \cdots s_{m+\sigma(m)}) \\ &\quad = s_0 + \sigma(0) + 1 + s_1 + \sigma(1) + 1 + \cdots + s_m + \sigma(m) + 1 \\ &\quad = 2 \cdot (0 + 1 + \cdots + m) + (m+1) \\ &\quad = m^2 + 2m + 1 \\ &\quad = (m+1)^2. \end{aligned}$$

We also observe that $s_i(0) = 0$ and thus $T \mid s_i$ for all $i \ge 0$. Therefore, $T^{m+1} \mid \det D$. Let $r \ge 1$ be an integer. Then

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

is the product of an $(m+1) \times r$ matrix and its transpose. So its rank is at most r and by Lemma 3.11 we have $(T-r)^{m+1-r} \mid \det D$ for all $r \in \{1, \ldots, m\}$. Next we use the known identity $B_i(T) = (-1)^i B_i(1-T)$ for all $i \ge 0$ and $B_i(0) = 0$ for all odd $i \ge 3$ to see that $s_i(1-T) = (-1)^{i+1} s_i(T)$ for all $i \ge 1$. For all $r \in \{2, \ldots, m+1\}$ we see that

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

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

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

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 the rank of $D(1-r)$ is 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 also holds 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 $s_i = \frac{T^{i+j+1}}{i+j+1}$+(lower order terms in T) and

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

Each term in this sum is $\frac{\operatorname{sign}(\sigma)}{i+j+1}$ + (terms of order <0 in T). We conclude that λ is the determinant of the $(m+1)\times(m+1)$ Hilbert matrix $\left(\frac{1}{i+j+1}\right)_{0\leq i,j\leq m}$. This determinant has been computed in [10] and is equal to $\lambda = G(m+2)^4/G(2m+3)$. This yields

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

Finally, using the definition of $\gamma_{N,m}$ (Definition 3.3) and (3.12) the lemma follows. □

3.3 Proof of Proposition 3.4

Proof. For part (i) observe that $B(1,\dots,1)$ is equal to the sum of the coefficients in $B(\overline{z_0},\dots,\overline{z_m},z_0,\dots z_m)$ and that each term can be bounded by $\prod_{j=0}^{m}\max\{1,|z_j|\}^{\deg_{z_j}B}$. Also note that $|z_j| = |\overline{z_j}|$, so using Lemma 3.12 we have $2\cdot\max_{0\leq j\leq m}\{\deg_{z_j}B, \deg_{\overline{z_j}}B\} \leq 2(N-(m+1))$. The above observations and the triangle inequality yield

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

So we find

$$
\begin{aligned}
|2X| &= |2\frac{1}{2}\log A_n(\overline{z_0},\dots,\overline{z_m})^\top A_N(z_0,\dots z_m)| \\
&= |\log B(\overline{z_0},\dots,\overline{z_m},z_0,\dots z_m)| + \left|\log \prod_{0\le i<j\le m}(\overline{z_j}-\overline{z_i})(z_j-z_i)\right| \\
&\le \log B(1,\dots,1)\prod_{j=0}^{m}\max\{1,|z_j|\}^{2(N-(m+1))} + \left|\log \prod_{0\le i<j\le m}(\overline{z_j}-\overline{z_i})(z_j-z_i)\right| \\
&\le \log \gamma_{N,m}N^{m+1} + \log\prod_{j=0}^{m}\max\{1,|z_j|\}^{2(N-(m+1))} + \log \prod_{0\le i<j\le m}|\overline{z_j}-\overline{z_i}||z_j-z_i| \\
&= \log \gamma_{N,m}N^{m+1} + 2(M-(m+1))\sum_{j=0}^{m}\log^+|z_j| + \sum_{0\le i<j\le m}\log|z_j-z_i|^2 \\
&= \log \gamma_{N,m}N^{m+1} + 2\sum_{0\le i<j\le m}\log|z_j-z_i| + 2(M-(m+1))\sum_{j=0}^{m}\log^+|z_j|
\end{aligned}
$$

and dividing by 2 yields part (i).

For part (ii) we use Hadamard's Inequality (Theorem A.1). 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 $\mathbb{C}^N$. Denote by $A'_N(z_0,\dots,z_m) = (a'_{i,j})$ the $N\times N$ matrix obtained by adding these columns to $A_N(z_0,\dots,z_m)$. Then we apply Hadamard's Inequality to find

$$
\begin{aligned}
|\det A'_N(z_0,\dots,z_m)| &\le \prod_{j=0}^{N-1}\left(\sum_{i=0}^{N-1}|a'_{i,j}|^2\right)^{1/2} \\
&= \prod_{j=0}^{m}\left(\sum_{i=0}^{N-1}|a'_{i,j}|^2\right)^{1/2}\cdot 1\cdot 1\cdots 1 \\
&= \prod_{j=0}^{m}(1+|z_j|^2+\cdots+|z_j|^{2(N-1)})^{1/2}.
\end{aligned}
$$

Similarly,

$$\begin{aligned}
|\det A'_N(\overline{z_0},\ldots,\overline{z_m})^\top| &= |\det A'_N(\overline{z_0},\ldots,\overline{z_m})| \\
&\le \prod_{j=0}^{N-1}\left(\sum_{i=0}^{N-1}|\overline{a'_{i,j}}|^2\right)^{1/2} \\
&= \prod_{j=0}^{m}\left(\sum_{i=0}^{N-1}|\overline{a'_{i,j}}|^2\right)^{1/2}\cdot 1\cdot 1\cdots 1 \\
&= \prod_{j=0}^{m}(1+|\overline{z_j}|^2+\cdots+|\overline{z_j}|^{2(N-1)})^{1/2} \\
&= \prod_{j=0}^{m}(1+|z_j|^2+\cdots+|z_j|^{2(N-1)})^{1/2}.
\end{aligned}$$

Now observe that because the matrix $A'_N(\overline{z_0},\ldots,\overline{z_m})^\top\cdot A'_N(z_0,\ldots,z_m)$ is of the form $\begin{pmatrix} A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m) & 0 \\ 0 & I_{N-m+1}\end{pmatrix}$, we have that

$$\det A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m) = \det A'_N(\overline{z_0},\ldots,\overline{z_m})^\top A'_N(z_0,\ldots,z_m).$$

So we can bound X by

$$\begin{aligned}
X &= \frac{1}{2}\log\left(\det A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m)\right) \\
&= \frac{1}{2}\log\left(\det A'_N(\overline{z_0},\ldots,\overline{z_m})^\top A'_N(z_0,\ldots,z_m)\right) \\
&= \frac{1}{2}\log\left(\det A'_N(\overline{z_0},\ldots,\overline{z_m})^\top\cdot\det A'_N(z_0,\ldots,z_m)\right) \\
&\le \frac{1}{2}\log\prod_{j=0}^{m}(1+|z_j|^2+\cdots+|z_j|^{2(N-1)}) \\
&= \log\prod_{j=0}^{m}(1+|z_j|^2+\cdots+|z_j|^{2(N-1)})^{1/2}.
\end{aligned}$$

Now,

$$
\begin{aligned}
X &\le \log \prod_{j=0}^{m} (1 + |z_j|^2 + \cdots + |z_j|^{2(N-1)})^{1/2} \\
&= \sum_{j=0}^{m} \log(1 + |z_j|^2 + \cdots + |z_j|^{2(N-1)})^{1/2} \\
&= \frac{1}{2} \sum_{j=0}^{m} \log(1 + |z_j|^2 + \cdots + |z_j|^{2(N-1)}) \\
&\le \frac{1}{2} \sum_{j=0}^{m} \log\left(1 + \max\{1, |z_j|\}^2 + \cdots + \max\{|1, z_j|\}^{2(N-1)}\right) \\
&\le \frac{1}{2} \sum_{j=0}^{m} \log\left(1^{2(N-1)} + \max\{1, |z_j|\}^{2(N-1)} + \cdots + \max\{1, |z_j|\}^{2(N-1)}\right) \\
&\le \frac{1}{2} \sum_{j=0}^{m} \log\left(N \max\{1, |z_j|\}^{2(N-1)}\right) \\
&= \frac{1}{2} \sum_{j=0}^{m} \log N + \frac{1}{2} \sum_{j=0}^{m} 2(N-1) \log^+(z_j) \\
&= \frac{m+1}{2} \log N + \sum_{j=0}^{m} (N-1) \log^+(z_j).
\end{aligned}
$$

Therefore we have

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

as desired.

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

$$\begin{aligned}
e^{2X} &= e^{2\frac{1}{2}\log\det A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m)} \\
&= \det A_N(\overline{z_0},\ldots,\overline{z_m})^\top A_N(z_0,\ldots,z_m) \\
&= \sum_I \det A_I(\overline{z_0},\ldots,\overline{z_m})^\top A_I(z_0,\ldots,z_m) \\
&\le \sum_I |\det A_I(z_0,\ldots,z_m)|^2 \\
&\le \binom{N}{m+1} \max_I |\det A_I(z_0,\ldots,z_m)|^2.
\end{aligned}$$

Moreover,

$$\begin{aligned}
|\det A_I(z_0,\ldots,z_m)| &\le (m+1)! \max_\sigma |z_0|^{\alpha_{\sigma(0)}} \cdots |z_m|^{\alpha_{\sigma(m)}} \\
&\le (m+1)! \max_\sigma \left\{ (\max\{|z_0|,\ldots,|z_m|\})^{\alpha_{\sigma(0)}+\cdots+\alpha_{\sigma(m)}} \right\},
\end{aligned}$$

where σ runs over all permutations of $\{0,\ldots,m\}$. We observe that

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

By the hypothesis of (iii) we have $|z_j| \le 1$ for all j, thus

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

Since $\binom{N}{m+1}^{1/2}(m+1)! \le \binom{N}{m+1}(m+1)! \le N^{m+1}$, we conclude that

$$\begin{aligned}
e^X &\le \binom{N}{m+1}^{1/2} (m+1)! \max\{|z_0|,\ldots,|z_m|\}^{\frac{m(m+1)}{2}} \\
&\le N^{m+1} \max\{|z_0|,\ldots,|z_m|\}^{\frac{m(m+1)}{2}},
\end{aligned}$$

and thus

$$\begin{aligned}
\log e^X = X &\le \log N^{m+1} + \log \max\{|z_0|,\ldots,|z_m|\}^{\frac{m(m+1)}{2}} \\
&= (m+1)\log N + \frac{m(m+1)}{2} \log \max\{|z_0,\ldots,z_m|\},
\end{aligned}$$

as desired. □

3.4 Asymptotic upper bound for $\gamma_{N,m}$

We start by defining

$$\chi(x) = -x \log x - x \log 4 + \frac{1+x^2}{2x} \log(1-x^2) + \log \frac{1+x}{1-x}, \tag{3.13}$$

for $x \in (0,1)$.

In this subsection we will give the following asymptotic upper bound for $\gamma_{N,m}$.

Lemma 3.13. *Let $N \geq 2$ and $m \geq 0$ be integers with $N \geq m+1$. Set $p = \frac{m+1}{N}$, then*

$$\frac{1}{2} \log \gamma_{N,m} \leq p\chi(p) \frac{N(N-1)}{2} + O((m+1)\log N). \tag{3.14}$$

The function χ extends to a continuous function on $[0,1]$ with $\chi(0) = \chi(1) = 0$. For $x \in (0,1)$ we have

$$\chi''(x) = \frac{\log(1-x^2)}{x^3} < 0. \tag{3.15}$$

So χ is concave on $[0,1]$ and in particular it takes non-negative values. Moreover, using that the Taylor series of $x \mapsto \log(1-x^2)$ equals to $-\sum_{k=1}^{\infty} \frac{x^{2k}}{k}$ we find

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

on $(0,1)$. Now using that $\chi(0) = 0$ and that $\lim_{x\to 0} = \frac{\chi(x)+x\log(x)}{x} = \frac{3}{2} - \log(4)$ we obtain

$$\chi(x) = -x \log x + \big(\frac{3}{2} - \log 4\big)x - \sum_{k=2}^{\infty} \frac{x^{2k-1}}{k(2k-2)(2k-1)} \tag{3.16}$$

on $(0,1)$, where we used the following

$$\begin{aligned}
&\frac{1+x^2}{2x}\log(1+x^2)+\log\left(\frac{1+x}{1-x}\right)\\
&=\frac{1+x^2}{2x}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}((-x)^2)^n+\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^n-\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}(-x)^n\\
&=\frac{1+x^2}{2x}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^{2n}+2\sum_{m=0}^{\infty}\frac{1}{2m+1}x^{2m+1}\quad(n=2m+1)\\
&=-\frac{1+x^2}{2x}\sum_{n=1}^{\infty}\frac{x^{2n}}{n}+2\sum_{n=0}^{\infty}\frac{1}{2n+1}x^{2n+1}\\
&=-\frac{1}{2x}\sum_{n=1}^{\infty}\frac{x^{2n}}{n}-\frac{1}{2}x\sum_{n=1}^{\infty}\frac{x^{2n}}{n}+2\sum_{n=0}^{\infty}\frac{1}{2n+1}x^{2n+1}\\
&=-\frac{1}{2}\sum_{n=1}^{\infty}\frac{x^{2n-1}}{n}-\frac{1}{2}\sum_{n=1}^{\infty}\frac{x^{2n+1}}{n}+2\sum_{n=0}^{\infty}\frac{1}{2n+1}x^{2n+1}\\
&=-\frac{1}{2}\sum_{n=0}^{\infty}\frac{x^{2n+1}}{n}-\frac{1}{2}\sum_{n=1}^{\infty}\frac{x^{2n+1}}{n}+2\sum_{n=0}^{\infty}\frac{1}{2n+1}x^{2n+1}\\
&=\frac{3}{2}x+\sum_{n=1}^{\infty}\left(-\frac{1}{2n+2}-\frac{1}{2n}+\frac{2}{2n+1}\right)x^{2n+1}\\
&=\frac{3}{2}+\sum_{n=1}^{\infty}\frac{-2n(2n+1)-(2n+1)(2n+2)+2\cdot 2n(2n+1)}{2n(2n+1)(2n+2)}x^{2n+1}\\
&=\frac{3}{2}+\sum_{n=1}^{\infty}\frac{x^{2n+1}}{2n(2n+1)(n+1)}\\
&=\frac{3}{2}+\sum_{n=1}^{\infty}\frac{x^{2k-1}}{(2k-2)k(2k-1)}\quad(n=k-1).
\end{aligned}$$

We now prove some properties of the Barnes G-function of Definition 3.2.

Lemma 3.14. *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 $f : [a,b] \to \mathbb{R}$ be a non decreasing continuous function. We will use the inequality

$$f(a) + \int_a^{b-1} f(x)dx \leq \sum_{j=a}^{b-1} f(j) \leq \int_a^b f(x)dx.$$

The map $x \mapsto x \log x - x$ is non decreasing on $x \geq 1$. So

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

We compute the two integrals using that $x \mapsto \frac{x^2}{2} \log x - \frac{3}{4}x^2$ is an anti derivative of $x \mapsto x \log x - x$ to find

$$\begin{aligned}
-1+&\int_1^{m-1} (x \log x - x)dx \\
&= -1 + \left[\frac{x^2}{2} \log x - \frac{3}{4}x^2\right]_1^{m-1} \\
&= -1 + \frac{(m-1)^2}{2} \log(m-1) - \frac{3}{4}(m-1)^2 + \frac{3}{4} \\
&= -1 + \frac{m^2}{2} \log(m-1) - m \log(m-1) + \frac{1}{2} \log(m-1) - \frac{3}{4}m^2 + \frac{3}{2}m \\
&= \frac{m^2}{2} \log(m) - \frac{3}{4}m^2 + O(m \log(m+1))
\end{aligned}$$

and

$$\begin{aligned}
\int_1^m (x \log x - x)dx &= \left[\frac{x^2}{2} \log x - \frac{3}{4}x^2\right]_1^m \\
&= \frac{m^2}{2} \log m - \frac{3}{4}m^2 + \frac{3}{4}.
\end{aligned}$$

Stirling's approximation (Proposition A.21) states that $\log j! = j \log j - j + O(\log(j+1))$. Thus

$$\begin{aligned}
\sum_{j=1}^{m-1} \log j! &= \sum_{j=1}^{m-1} j \log j - j + \sum_{j=1}^{m-1} O(\log(j+1)) \\
&= \sum_{j=1}^{m-1} j \log j - j + O(m \log(m+1)).
\end{aligned}$$

Putting everything together proves the lemma. □

Lemma 3.15. *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 3.14 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 $2 \log G(m+2) - G(2m+3)$. ☐

Lemma 3.16. *Let N and $m \geq 0$ be integers with $N \geq m+2$. Set $p = (m+1)/N$, then*

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

Proof. Let a and b be integers with $a \leq b$ and suppose $f : [a,b] \mapsto \mathbb{R}$ 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

$$\begin{aligned} S &\leq 2(m+1)\log N + \int_0^{m+1} (m+1-x)(\log(N-x) + \log(N+x))dx \\ &= 2(m+1)\log N + 2\log(N) \int_0^{m+1} (m+1-x)dx + N^2 \int_0^p (p-y)\log(1-y^2)dy \end{aligned} \tag{3.17}$$

after a substitutionn $y = x/N$. 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 the definition of χ. This allows us to compute the integral in (3.17) and conclude the proof. □

Proof of Lemma 3.13. If $n=m+1$, then the lemma follows from $\chi(1)=0$ and (3.10). So we may assume $N\geq m+2$. Using the definition of $\gamma_{N,m}$ (Definition 3.3), we write

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

Adding the bounds from Lemmas 3.15 and 3.16 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 (3.14); indeed, the continuous function $\chi:[0,1]\to\mathbb{R}$ is bounded from above. □

4 Transfinite diameter of a hedgehog

In this section we bound from above the transfinite diameter of a hedgehog following the work of Habegger [9].

4.1 Preliminary results

Lemma 4.1. *For any real number $z \in [-2, 2]$ there is a $w \in \mathbb{C} \setminus \{0\}$ with $|w| = 1$ such that $w + w^{-1} = z$.*

Proof. We write $w = e^{i\theta} = \cos(\theta) + i\sin(\theta)$ for some real number θ. We compute the sum $w + w^{-1}$

$$w + w^{-1} = (\cos(\theta) + i\sin(\theta)) + (\cos(\theta) - i\sin(\theta)) = 2\cos(\theta).$$

We need to show that

$$w + w^{-1} = 2\cos(\theta) = z$$

or, equivalently, that

$$\cos(\theta) = \frac{z}{2}$$

Since $z \in [-2, 2]$, it follows that $\frac{z}{2} \in [-1, 1]$. The cosine function maps real numbers to the interval $[-1, 1]$, so there exists a real number ζ such that

$$\cos(\zeta) = \frac{z}{2}.$$

Let $w = e^{i\zeta}$. By construction, $|w| = 1$ and it satisfies

$$w + w^{-1} = 2\cos(\zeta) = z.$$

□

Lemma 4.2. *Let $m \geq 0$ and suppose $z_0, \ldots, z_m$ lie on a line segment of length ϵ. Then*

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

Proof. The left-hand side of (4.1) is invariant under translating all z_i. If we translate appropriately and stretch by the factor $\frac{4}{\epsilon}$, then

$$\prod_{0\le i<j\le m} \left(\frac{4}{\epsilon}\right)|z_j - z_i| = \left(\frac{4}{\epsilon}\right)^{\frac{m(m+1)}{2}} \prod_{0\le i<j\le m} |z_j - z_i|,$$

so the product is multiplied by $\left(\frac{4}{\epsilon}\right)^{m(m+1)/2}$. Without loss of generality, we may thus assume $\epsilon = 4$, that the line segment in question equals $[-2, 2]$, and that $m \ge 1$. For each i we have $z_i \in [-2, 2]$, hence by Lemma 4.1 there is $w_i \in \mathbb{C} \setminus \{0\}$ with $|w_i| = 1$ and $w_i + w_i^{-1} = z_i$. Let

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

We now apply Lemma 2 in [11]. In our case, using $|w_j| = 1$, it implies that

$$\begin{aligned} V &= \frac{1}{2}\left|\det\left(w_j^i + w_j^{-i}\right)_{0\le i,j\le m}\right| \\ &= \frac{1}{2}\begin{vmatrix} w_0^0 + w_0^{-0} & w_1^0 + w_1^{-0} & w_2^0 + w_2^{-0} & \cdots & w_m^0 + w_m^{-0} \\ w_0^1 + w_0^{-1} & w_1^1 + w_1^{-1} & w_2^1 + w_2^{-1} & \cdots & w_m^1 + w_m^{-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_0^m + w_0^{-m} & w_1^m + w_1^{-m} & w_2^m + w_2^{-m} & \dots & w_m^m + w_m^{-m} \end{vmatrix}. \end{aligned}$$

We have $|w_j^i + w_j^{-i}| \le 2$ for all i and j. Hadamard's inequality (Theorem A.1) implies

$$\begin{aligned} V &\le \frac{1}{2}\prod_{j=0}^{m}\left(\sum_{i=0}^{m} |w_j^i + w_j^{-i}|^2\right)^{1/2} \\ &\le \frac{1}{2}\prod_{j=0}^{m}\left((m+1)\cdot 2^2\right)^{1/2} \\ &\le \frac{1}{2}2^{m+1}(m+1)^{(m+1)/2} \\ &= 2^m(m+1)^{(m+1)/2}. \end{aligned}$$

□

Lemma 4.3 (Fischer's Inequality)**.** *Let $n \in \mathbb{N}$, let $m_1, \ldots, m_n \geq 0$ be integers, and set $N = (m_1+1) + \cdots + (m_n+1)$. For each $l \in \{1, \ldots, n\}$ let $M_l \in Mat_{N,m_l+1}(\mathbb{C})$ and set $M = (M_1 \cdots M_n) \in \mathrm{Mat}_N(\mathbb{C})$. Then*

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

Proof. Let $\begin{pmatrix} M' & * \\ * & M'' \end{pmatrix} \in \mathrm{Mat}_N(\mathbb{C})$ be a positive definite Hermitian matrix with $M' \in \mathrm{Mat}_r(\mathbb{C})$ and $M'' \in \mathrm{Mat}_{N-r}(\mathbb{C})$. Theorem 13.5.5 [17] states that

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

If the $N \times N$ matrix is merely positive semi-definite, then adding a positive multiple of the unit matrix leads to a positive-definite Hermitian matrix. So the inequality holds for all positive semi-definite matrices by continuity. Moreover, by induction the analog inequality holds for more than two matrices on the diagonal.

As $\overline{M}^\top M$ is positive semidefinite and Hermitian we conclude

$$\det \overline{M}^\top M = \det \begin{pmatrix} \overline{M_1}^\top \\ \vdots \\ \overline{M_n}^\top \end{pmatrix} \begin{pmatrix} \overline{M_1}^\top & \cdots & \overline{M_n}^\top \end{pmatrix}$$

$$= \det \begin{pmatrix} \overline{M_1}^\top M_1 & & * \\ & \ddots & \\ * & & \overline{M_n}^\top M_n \end{pmatrix} \leq \prod_{l=1}^{n} \det \overline{M_l}^\top M_l.$$

□

4.2 Habegger's bound

The following result due to Habegger [9] bounds from above the transfinite diameter of the union of a hedgehog with n quils and a closed disk of radius $1 - 1/n$ centered at the origin.

Theorem 4.4. *Let $n \in \mathbb{Z}_{>0}$ and $a_1, \ldots, a_n \in \mathbb{C}$ with $\max\{|a_1|, \ldots, |a_n|\} = 1$. Set $\mathcal{K} = \mathcal{K}(a_1, \ldots, a_n)$ and $\mathcal{S} = \mathcal{K} \cup \{z \in \mathbb{C} : |z| \le 1 - 1/n\}$. Then*

$$\log \tau_N(\mathcal{K}) \le \log \tau_N(\mathcal{S}) \le -\frac{0.39}{n} + O\left(\frac{\log(nN)}{N}\right).$$

In particular, $\tau(\mathcal{K}) \le \tau(\mathcal{S}) \le e^{-0.39/n}$.

Corollary 4.5. *Let $n \in \mathbb{Z}_{>0}$ and $a_1, \ldots, a_n \in \mathbb{C}$ and set $\mathcal{K} = \mathcal{K}(a_1, \ldots, a_n)$. Then*

$$\tau(\mathcal{K}) \le e^{-0.39/n} \max_{1 \le i \le n} |a_i|.$$

Proof. Define

$$\mathcal{K}' = \mathcal{K}\left(\frac{a_1}{\max_{1\le i\le n}|a_i|}, \ldots, \frac{a_d}{\max_{1\le i\le n}|a_i|}\right).$$

Then by Theorem 4.4 $\tau(\mathcal{K}') \le e^{-0.39/n}$. Now since $\mathcal{K} = \mathcal{K}' \max_{1\le i\le n}|a_i|$ we obtain $\tau(\mathcal{K}) = \tau(\mathcal{K}')\max_{1\le i\le n}|a_i| \le e^{-0.39/2d}\max_{1\le i\le n}|a_i|$ by the scaling property of the transfinite diameter. □

Proof of Theorem 4.4. The first inequality follows as $\tau_N(K) \le \tau_N(L)$ for all non-empty compact subsets $K \subset L \subset \mathbb{C}$ and all $N \ge 2$. We now prove the bound for $\mathcal{S}$.

Let $\epsilon = 1/n \in (0, 1]$. Our choice of ϵ is in part made by convenience. Let $N \ge 2$ and suppose $z_1, \ldots, z_N \in \mathcal{S}$ are pairwise distinct. Our goal is to bound

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

from above where the second equality follows as the matrix is of Vandermonde type (Proposition A.3).

We arrange our points $z_1, \ldots, z_N$ into $n+1$ parts as follows. We first collect all points z_j with $|z_j| \le 1 - \epsilon$, relabel these points $z_{0,0}, \ldots, z_{0,m_0}$. If $|z_j| > 1 - \epsilon$, fix any $l \in \{1, \ldots, n\}$ with $z_j \in [0, 1]a_l$. We add z_j to the l-th part. So for each $l \in \{1, \ldots, n\}$ we obtain points $z_{l,0}, \ldots, z_{l,m_l}$ on $[0, 1]a_l$. Note that $(m_0 + 1) + \cdot + (m_n + 1) = N$ and $m_l \ge -1$ for all l. We set $p_l = (m_l + 1)/N$.

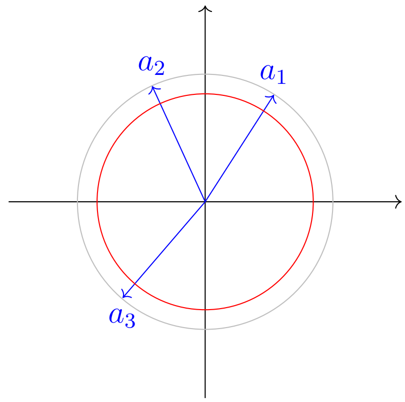

Figure 4: The inner circle has radius $1-1/n$ and the outer circle has radius 1. The points $z_{0,0},\dots,z_{0,m_0}$ are inside the red circle. For each $l\in\{1,\dots,n\}$ the points $z_{l,0},\dots,z_{l,m_l}$ are outside the inner circle and on the line segment $[0,1]a_l$.

Fischer's Inequality (Lemma 4.3), implies

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

where $A_N(z_{l,0},\dots,z_{l,m_l})\in \mathrm{Mat}_{N,m+1}(\mathbb{C})$ as in (3.1).

Recall that $|z_{l,j}|\le 1$. We will apply Corollary 3.5 to the terms $l\in\{1,\dots,n\}$ and Proposition 3.4 part (iii) to $l=0$, if $m_l\ge 0$, respectively to obtain the following bound:

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

We now prove (4.2). We first bound the term of the sum for $l=0$ if $m_0\ge 0$

using Proposition 3.4 (iii):

$$
\begin{aligned}
&\frac{2}{N(N-1)}\frac{1}{2}\log|\det A_N(\overline{z_{0,0}},\dots\overline{z_{0,m_0}})^\top A_N(z_{0,0},\dots,z_{0,m_0})| \\
&\le \frac{2}{N(N-1)}\left((m_0+1)\log N + \frac{m_0(m_0+1)}{2}\log\underbrace{\max\{|z_{0,0}|,\dots,|z_{0,m_0}|\}}_{\le 1-\epsilon}\right) \\
&\le 2\frac{m_0+1}{N-1}\frac{\log N}{N} + \frac{m_0+1}{N}\frac{m_0}{N-1}\log(1-\epsilon) \\
&\le O\left(\frac{\log N}{N}\right) + p_0\frac{m_0}{N-1}\log(1-\epsilon).
\end{aligned} \tag{4.3}
$$

Now using Corollary 3.5 to bound each term of the sum for $1 \le l \le n, m_l \ge 0$ we obtain

$$
\begin{aligned}
&\frac{2}{N(N-1)}\sum_{\substack{l=1\\ m_l\ge 0}}^{n}\frac{1}{2}\log|\det A_N(\overline{z_{l,0}},\dots\overline{z_{l,m_l}})^\top A_N(z_{l,0},\dots,z_{l,m_l})| \\
&\le \frac{2}{N(N-1)}\sum_{\substack{l=1\\ m_l\ge 0}}^{n}\left(\min\left\{0,\frac{1}{2}\log\gamma_{N,m} + \sum_{0\le i,j\le m}\log|z_j-z_i|\right.\right. \\
&\left.\left. - m\sum_{j=0}^{m}\log^+|z_j|\right\} + \frac{m+1}{2}\log N + (N+1))\sum_{j=0}^{m}\log^+|z_j|\right) \\
&\le \sum_{\substack{l=1\\ m_l\ge 0}}^{n}\min\left\{0, p_l\chi(p_l) + \frac{2}{N(N-1)}\sum_{0\le i,j\le m_l}\log|z_j-z_i|\right\} \\
&\qquad + O\left(\frac{(m_1+1)+\cdots+(m_n+1)}{(N-1)}\frac{\log N}{N}\right) \\
&= \sum_{\substack{l=1\\ m_l\ge 0}}^{n}\min\left\{0, p_l\chi(p_l) + \frac{2}{N(N-1)}\sum_{0\le i,j\le m_l}\log|z_j-z_i|\right\} \\
&\qquad + O\left(\frac{\log N}{N}\right),
\end{aligned} \tag{4.4}
$$

where we used Lemma 3.13 for the second inequality, we used $(m_0+1)+\cdots+(m_n+1) = N$ to bound the error term, and observe that $\sum_{j=0}^{m}\log^+|z_j| = 0$

because $|z_{l,j}| \leq 1$. Note also that every term coming from some l with $m_l \leq -1$ can be omitted. Combining the two cases (4.3) and (4.4) together yield (4.2).

Let us treat the terms on the right-hand side separately. For $l = 0$ and if $m_0 \geq 0$ we use $\frac{m_0}{N-1} = p_0 - \frac{N-(m_0+1)}{N(N-1)}$ and $\log(1-\epsilon) \leq -\epsilon$ to find

$$p_0 \frac{m_0}{N-1} \log(1-\epsilon) \leq -\epsilon p_0^2 + \epsilon p_0 \frac{N-(m_0+1)}{N(N-1)} = -\epsilon p_0^2 + O\left(\frac{1}{N}\right). \tag{4.5}$$

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

$$\begin{aligned}
&\sum_{0\leq i<j\leq m_l} \log|z_{l,j} - z_{l,i}| \\
&\qquad \leq \log\left(2^{m_l}(m_l+1)^{(m_l+1)/2}\left(\frac{\epsilon}{4}\right)^{m_l(m_l+1)/2}\right) \\
&\qquad = m_l \log(2) + \frac{m_l+1}{2}\log(m_l+1) + \frac{m_l(m_l+1)}{2}\log\left(\frac{\epsilon}{4}\right) \\
&\qquad = \frac{m_l(m_l+1)}{2}\log\frac{\epsilon}{4} + O((m_l+1)\log(m_l+1)).
\end{aligned} \tag{4.6}$$

Recall that $\frac{m_l}{N-1} = p_l - \frac{N-(m+1)}{N(N-1)}$ and $m_l + 1 \leq N$ and note that

$$m_l = (N-1)p_l - \frac{N-(m+1)}{N} \leq (N-1)p_l \tag{4.7}$$

and

$$m_l + 1 \leq (N-1)p_l + 1. \tag{4.8}$$

We substitute (4.7) and (4.8) into (4.6) to obtain

$$\begin{aligned}
&\sum_{0\leq i<j\leq m_l} \log|z_{l,j} - z_{l,i}| \\
&\leq \frac{(N-1)p_l\,((N-1)p_l+1)}{2}\log\left(\frac{\epsilon}{4}\right) + O\left((N-1)p_l\log\left((N-1)p_l+1\right)\right).
\end{aligned}$$

We multiply by $\frac{2}{N(N-1)}$

$$
\begin{aligned}
&\frac{2}{N(N-1)} \sum_{0\le i<j\le m_l} \log|z_{l,j}-z_{l,i}| \\
&\le \frac{2}{N(N-1)} \frac{(N-1)^2 p_l^2 + (N-1)p_l}{2} \log\left(\frac{\epsilon}{4}\right) \\
&\qquad\qquad\qquad\qquad + O\left(\frac{2}{N(N-1)}(N-1)p_l\left((N-1)p_l+1\right)\right) \\
&= \frac{(N-1)p_l^2 + p_l}{N} \log\left(\frac{\epsilon}{4}\right) + O\left(\frac{2}{N} p_l \log\left((N-1)p_l+1\right)\right) \\
&\le p_l^2 \log\left(\frac{\epsilon}{4}\right) + p_l \log\left(\frac{\epsilon}{4}\right) + O\left(\frac{1}{N} p_l \log\left((N-1)p_l+1\right)\right) \\
&\le p_l^2 \log\left(\frac{\epsilon}{4}\right) + O\left(\frac{1}{N} p_l \log\left((N-1)p_l+1\right)\right) \quad \left(\text{because } \log\left(\frac{\epsilon}{4}\right) < 0\right) \\
&= p_l^2 \log\left(\frac{\epsilon}{4}\right) + O\left(\frac{1}{N} p_l \log\left(N p_l\right)\right) \\
&\le p_l^2 \log\left(\frac{\epsilon}{4}\right) + O\left(\frac{1}{N} p_l \log\left(N\right)\right) \quad (\text{because } p_l \le 1).
\end{aligned}
$$

So using that $\log(N) \le \log(4N/\epsilon)$ we get

$$
\frac{2}{N(N-1)} \sum_{0\le i<j\le m_l} \log|z_{l,j}-z_{l,i}| \le p_l^2 \log\frac{\epsilon}{4} + O\left(p_l \frac{\log(4N/\epsilon)}{N}\right). \qquad (4.9)
$$

We plug (4.5) and (4.9) into (4.2) and find

$$
\begin{aligned}
v &\le -\epsilon p_0^2 + O\left(\frac{1}{N}\right)\sum_{l=1}^{n}\min\left\{0, p_l\chi(p_l) + p_l^2\log\frac{\epsilon}{4} + O\left(p_l\frac{\log(4N/\epsilon)}{N}\right)\right\} \\
&\le -\epsilon p_0^2 + O\left(\frac{1}{N}\right)\sum_{l=1}^{n} p_l O\left(\frac{\log(4N/\epsilon)}{N}\right) + p_l\min\left\{0, \chi(p_l) + p_l\log\frac{\epsilon}{4}\right\} \\
&\le -\epsilon p_0^2 + O\left(\frac{1}{N}\right) + (p_1 + \cdots + p_n)O\left(\frac{\log(4N/\epsilon)}{N}\right) \\
&\qquad + \sum_{l=1}^{n} p_l\min\left\{0, \chi(p_l) + p_l\log\frac{\epsilon}{4}\right\} \\
&\le -\epsilon p_0^2 + \sum_{l=1}^{n} p_l\min\left\{0, \chi(p_l) + p_l\log\frac{\epsilon}{4}\right\} + O\left(\frac{\log(4N/\epsilon)}{N}\right)
\end{aligned}
$$

as $p_1+\cdots+p_n \le 1$. Next recall that $x \mapsto \chi(x)$ is concave on $[0,1]$. 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 (Proposition A.22) implies

$$
\begin{aligned}
&\sum_{l=1}^{n} p_l\min\left\{0, \chi(p_l) + p_l\log\frac{\epsilon}{4}\right\} = \left(\sum_{l=1}^{n} p_l\right)\cdot\frac{\sum_{l=1}^{n} p_l\min\left\{0, \chi(p_l) + p_l\log\frac{\epsilon}{4}\right\}}{\sum_{l=1}^{n} p_l} \\
&\le \left(\sum_{l=1}^{n} p_l\right)\cdot\min\left\{0, \chi\left(\frac{p_1^2+\cdots+p_n^2}{p_1+\ldots p_n}\right) + \left(\frac{p_1^2+\cdots+p_n^2}{p_1+\ldots p_n}\right)\log\frac{\epsilon}{4}\right\}.
\end{aligned}
$$

And thus

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

with $p = p_1 + \cdots + p_n = 1 - p_0$ and $\sigma^2 = p_1^2 + \cdots + p_n^2$; if $p = 0$ then the bound holds when omitting the term $p\min\{\cdots\}$ here and corresponding terms below. By (3.16) we have

$$\begin{aligned}
\chi(x) &= -x\log x + \left(\frac{3}{2} - \log 4\right)x - \sum_{k=2}^{\infty} \frac{x^{2k-1}}{k(2k-2)(2k-1)} \\
&\leq -x\log x + \left(\frac{3}{2} - \log 4\right)x \\
&\leq -x\log x + \log e^{x\frac{3}{2} - x\log 4} \\
&= -\left(x\log x + x\log 4 - x\log e^{3/2}\right) \\
&= -x\log\left(\frac{4x}{e^{3/2}}\right)
\end{aligned}$$

for all $x \in (0,1]$. So

$$\begin{aligned}
p\left(\chi\left(\frac{\sigma^2}{p}\right) + \frac{\sigma^2}{p}\log\frac{\epsilon}{4}\right) &\leq p\frac{-\sigma^2}{p}\log\frac{\frac{4\sigma^2}{p}}{e^{3/2}} + p\frac{\sigma^2}{p}\log\frac{\epsilon}{4} \\
&= \sigma^2\left(\log\frac{4\sigma^4}{pe^{3/2}} - \log\frac{\epsilon}{4}\right) \\
&= -\sigma^2\log\frac{16\sigma^2}{e^{3/2}p\epsilon},
\end{aligned}$$

and thus

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

Lemma A.20 implies

$$\frac{p^2}{n} = \frac{(p_1 + \cdots + p_n)^2}{1 + \cdots + 1} \leq \frac{p_1^1}{1} + \cdots + \frac{p_n^2}{1} = \sigma^2$$

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

$$v \leq -\frac{1}{n}\left((1-p)^2 + p^2\log^+\left(\frac{16p}{e^{3/2}}\right)\right) + O\left(\frac{\log(nN)}{N}\right). \tag{4.10}$$

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

$$v \le -\frac{1}{n}\left(1-\frac{e^{3/2}}{16}\right)^2 + O\left(\frac{\log(nN)}{N}\right) \le \frac{1}{2n} + O\left(\frac{\log(nN)}{N}\right). \tag{4.11}$$

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

If $p \ge e^{3/2}/16$, then we can replace $\log^+$ by log in (4.10) and conclude

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

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

$$f(p) \ge f(p_0) \ge (1-0.488)^2 + 0.487^2 \log(16 \cdot 0.487 e^{-3/2}) > 0.39.$$

So

$$v \le -\frac{0.39}{n} + O\left(\frac{\log(nN)}{N}\right). \tag{4.12}$$

Regardless of the size of p we have (4.12) by (4.11).

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

$$\log \tau_N(\mathcal{S}) = \sup_{z_1,\ldots,z_N \in \mathcal{K}} \frac{2}{N(N-1)} \log \prod_{0 \le i < j \le N} |z_j - z_i| \le -\frac{0.39}{n} + O\left(\frac{log(nN)}{N}\right).$$

Taking the limit as $N \to \infty$ yields $\log \tau(\mathcal{S}) \le -0.39/n$, as desired. □

5 Proof of the Schinzel-Zassenhaus Conjecture

In this section, we prove Theorem 1.1. Before going through the proof, we need some preliminary results on the rationality of series.

5.1 Rationality results for series

Proposition 5.1. *A necessary and sufficient condition for a series*

$$f(z) = \sum_{j=0}^{\infty} f_j z^j \in \mathbb{C}[[z]]$$

to represent a rational function is that the determinants $\det \Delta_k$ *of the matrices*

$$\Delta_k = \begin{pmatrix} f_0 & f_1 & \cdots & f_k \\ f_1 & f_2 & \cdots & f_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ f_k & f_{k+1} & \dots & f_{2k} \end{pmatrix}, \tag{5.1}$$

are trivial for all sufficiently large k.

Proof. Suppose first that $f(z)$ is a rational function, $P(z)/Q(z)$ say, with $Q(z) = q_r + q_{r-1}z + \cdots + q_0 z^r$ and the q_j not all 0. Then from $Qf = P$, we see that

$$q_0 f_m + q_1 f_{m+1} + \cdots + q_r f_{m+r} = 0 \;\text{ for } m > \deg P. \tag{5.2}$$

This is the coefficient of the term that has degree $m + r > \deg P$ and thus it has to be zero.

Thus, for $k > r + \deg P$,

$$\Delta_k = \begin{pmatrix} f_0 & f_1 & \cdots & f_r & \cdots & f_{k-r} & \cdots & f_{r+\deg P} & f_{r+\deg P+1} \cdots & f_k \\ f_1 & f_2 & \cdots & f_{r+1} & \cdots & f_{k-r+1} & \cdots & f_{r+\deg P+1} & f_{r+\deg P+2} \cdots & f_{k+1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \quad\quad \ddots & \vdots \\ f_k & f_{k+1} & \dots & f_{k+r} & \cdots & f_{k-r+k} & \cdots & f_{k+r+\deg P} & f_{k+r+\deg P+1} \cdots & f_{2k} \end{pmatrix},$$

and the rightmost $r+1$ columns of Δ_k are linearly dependent, so $\det \Delta_k = 0$.

Conversely, suppose that $\det \Delta_k = 0$ for all $k \geq p$, where we can assume that p is the smallest integer with this property. Then the rightmost column of Δ_p is a linear combination of the first p columns of Δ_p, so that for some $q_0, q_1, \ldots, q_{p-1}$, we have

$$L_{p+j} := q_0 f_j + q_1 f_{j+1} + \cdots + q_{p-1} f_{j+p-1} + f_{p+j} = 0 \quad (j = 0, 1, \ldots, p).$$

We now show by induction that $L_{p+j} = 0$ for all $j \geq 0$. Assume that for some $m > p$ we have $L_{p+j} = 0$ for $j = 0, 1, 2, \ldots m-1$. We need to show that $L_{p+m} = 0$. Let us write Δ_m as

$$\Delta_m = \left(\begin{array}{cccc|cccc} & & & & f_p & f_{p+1} & \cdots & f_m \\ & \Delta_{p-1} & & & f_{p+1} & f_{p+2} & \cdots & f_{m+1} \\ & & & & \vdots & \vdots & \ddots & \vdots \\ & & & & f_{2p-1} & f_{2p} & \cdots & f_{p+m-1} \\ \hline f_p & f_{p+1} & \cdots & f_{2p-1} & f_{2p} & f_{2p+1} & \cdots & f_{p+m} \\ f_{p+1} & f_{p+2} & \cdots & f_{2p} & f_{2p+1} & f_{2p+2} & \cdots & f_{p+m+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_m & f_{m+1} & \cdots & f_{p+m-1} & f_{p+m} & f_{p+m+1} & \cdots & f_{2m} \end{array}\right).$$

Now, starting at the (p+1)th column, add to each column a linear combination of the previous p columns with coefficients $q_0, q_1, \ldots, q_{p-1}$. This gives

$$\begin{aligned} \det \Delta_m &= \det \left(\begin{array}{cccc|cccc} & & & & L_p & L_{p+1} & \cdots & L_m \\ & \Delta_{p-1} & & & \vdots & \vdots & \ddots & \vdots \\ & & & & L_{2p-1} & L_{2p} & \cdots & L_{p+m-1} \\ \hline f_p & f_{p+1} & \cdots & f_{2p-1} & L_{2p} & L_{2p+1} & \cdots & L_{p+m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f_m & f_{m+1} & \cdots & f_{p+m-1} & L_{p+m} & L_{p+m+1} & \cdots & L_{2m} \end{array}\right) \\ &= \det \left(\begin{array}{cccc|ccc} & \Delta_{p-1} & & & & 0 & \\ \hline f_p & f_{p+1} & \cdots & f_{2p-1} & 0 & & L_{p+m} \\ \vdots & \vdots & \ddots & \vdots & & \iddots & \vdots \\ f_m & f_{m+1} & \cdots & f_{p+m-1} & L_{p+m} & \cdots & L_{2m} \end{array}\right) \\ &= \pm (L_{p+m})^{m-p+1} \det \Delta_{p-1}. \end{aligned}$$

Then since $\det \Delta_m = 0$ and $\det \Delta_{p-1} \neq 0$, we have $L_{p+m} = 0$, completing the induction step.

Now we have L_{p+j} for $j > p$ equals to zero and

$$L_{p+j} = q_0 f_j + q_1 f_{j+1} + \cdots + q_{p-1} f_{j+p-1} + f_{p+j}$$

for $j > p$. So $f = P/Q$ for $\deg P \geq p-1, \deg Q = p$.
Thus, $f(z)$ represents a rational function with numerator of degree $p-1$ and denominator of degree p. □

In what follows we use the notation $\widehat{\mathbb{C}}$ for $\mathbb{C} \cup \{\infty\}$.

Theorem 5.2. *Let E be a compact subset of $\mathbb{C}$ that is symmetric about the real axis, so that $\overline{E} = E$, where $\overline{E}$ is the set of complex conjugates of elements of E. Suppose that the transfinite diameter $\tau(E)$ of E is less than 1. Suppose also that*

$$f\left(\frac{1}{z}\right) = \sum_{j=0}^{\infty} \frac{a_j}{z^j}$$

is regular on the complement $\widehat{\mathbb{C}} \setminus E$ of E and that all of the a_j are integers. Then $f(1/z)$ is a rational function.

Proof. Let $\epsilon > 0$, and E_ϵ be an ϵ-thickening of E. We can, by choosing ϵ sufficiently small, ensure that E_ϵ has also transfinite diameter, τ_ϵ say, less than 1. Note that $f(1/z)$ remains regular in $\widehat{\mathbb{C}} \setminus E_\epsilon$. Next, consider the $k \times k$ matrices

$$\Delta_k := \begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ a_2 & a_3 & \cdots & a_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_k & a_{k+1} & \dots & a_{2k} \end{pmatrix}$$

and $C_k := (c_{ij})_{i,j=0,1,\dots,k-1}$, where c_{ij} is the residue of $f(1/z)T_i(z)T_j(z)$ at $z = 0$. Here, the polynomials $T_i(z)$ of degree i are the Chebyshev polynomials for the set E_ϵ ; the T_i have maximum modulus m_i on E_ϵ (see Definition 2.2). Now define the $k \times k$ upper triangular matrix

$$B_k := (b_{ij})_{i,j=0,1,\dots,k-1} = \begin{pmatrix} 1 & b_{01} & \cdots & b_{0,k-1} \\ & 1 & & b_{1,k-1} \\ & & \ddots & \vdots \\ & 0 & & 1 \end{pmatrix},$$

where for any given j and $i \leq j$ its entries are defined by

$$T_j(z) = b_{jj} z^j + b_{j-1,j} z^{j-1} + \cdots + b_{0j}.$$

Note that as the polynomials T_j are all monic (so $b_{jj} = 1$), the matrix B_k has all its diagonal entries equal to 1. Then we have the identity

$$B_k^{\intercal} \Delta_k B_k = C,$$

because the matrices on both sides of the equation have (i, j)-th entry

$$\sum_{0 \le l \le i} \sum_{0 \le m \le j} b_{li} a_{l+m+1} b_{mj} \quad \text{for } 0 \le i \le k-1,\ 0 \le j \le k-1,$$

where $a_{l+m+1} = a_{l,m}$. Hence, $\det \Delta_k = \det C$. This identity will give us a way of bounding $|\det \Delta_k|$ from above.

We note that $f(1/z)$ is bounded on the boundary of E_ϵ, where ϵ has been chosen as above. Hence, by Cauchy's Integral Theorem, $|c_{ij}|$ is bounded above by a constant times $m_i m_j$. Specifically

$$|c_{i,j}| = \left| \frac{1}{2\pi i} \oint_\gamma f\left(\frac{1}{z}\right) T_i T_j \right| \le m_i m_j \left| \frac{1}{2\pi i} \oint_\gamma f\left(\frac{1}{z}\right) \right|.$$

The above holds because according to the residue theorem, for a meromorphic function f, the residue at the point a_k is given by:

$$\mathbf{Res}(f, a_k) = \frac{1}{2\pi i} \oint_c f(z) dz$$

where γ is a positively oriented simple closed curve around a_k and not including any other singularities on or inside the curve.

Choose a number τ_2 with $\tau_\epsilon < \tau_2 < 1$ and then an integer I such that for all $i \ge I$ we have $m_i^{1/i} < \tau_2$ or $m_i < \tau_2^i$. Next, we apply Hadamard's Inequality (Theorem A.1) to $\det C$. For $i < I$ there is a constant, c_1 say, with $\sum_{j=1}^{\infty} |c_{ij}|^2 \le c_1$, independent of k. This comes from

$$\begin{aligned} \sum_{j=1}^{\infty} |c_{ij}|^2 &\le \sum_{j=1}^{\infty} \underbrace{m_i^2}_{\text{constant}} m_j^2 \underbrace{\left| \frac{1}{2\pi i} \int_\gamma f\left(\frac{1}{z}\right) \right|^2}_{\text{constant}} \\ &\le \sum_{j=1}^{\infty} \tau_2^{2j} \underbrace{\max_{i \le I}\{m_i^2\}}_{\text{constant}} \underbrace{\left| \frac{1}{2\pi i} \int_\gamma f\left(\frac{1}{z}\right) \right|^2}_{\text{constant}} \le c_1 \quad (\text{where } i < I). \end{aligned}$$

The last inequality holds because for x such that $|x| < 1$ we have

$$\sum_{x=0}^{\infty} x^n = \frac{1}{1-x}.$$

While for $i \geq I$ there is a constant c_2 with $\sum_{j=1}^{\infty} |c_{ij}|^2 \leq c_2 \tau_2^{2i}$. This comes from

$$\begin{aligned}\sum_{j=1}^{\infty} |c_{ij}|^2 &= \sum_{j=1}^{\infty} m_i^2 m_j^2 \left| \frac{1}{2\pi i} \int_\gamma f\left(\frac{1}{z}\right) \right|^2 \\ &\leq \tau_2^{2i} \sum_{j=1}^{\infty} \tau_2^{2j} \underbrace{\left| \frac{1}{2\pi i} \int_\gamma f\left(\frac{1}{z}\right) \right|^2}_{\text{constant}} = \tau_2^{2i} c_2 \;\; (\text{where } i \geq I).\end{aligned}$$

Using Hadamard's (Theorem A.1) inequality we obtain

$$\begin{aligned}|\det \Delta_k|^2 = |\det C|^2 \leq \prod_{i=0}^{k-1} \left(\sum_{j=0}^{k-1} |c_{ij}|^2 \right) &= \prod_{i=0}^{I-1} \underbrace{\left(\sum_{j=0}^{k-1} |c_{ij}|^2 \right)}_{\leq c_1} \cdot \prod_{i=I}^{k-1} \underbrace{\left(\sum_{j=0}^{k-1} |c_{ij}|^2 \right)}_{\leq \tau_2^{2i} c_2} \\ &\leq c_1^I \cdot \prod_{i=I}^{k-1} \tau_2^{2i} c_2 = c_3 c_2^{k-I} \prod_{i=I}^{k-1} \tau_2^{2i}\end{aligned}$$

Now we look at the product

$$\begin{aligned}\prod_{i=I}^{k-1} \tau_2^{2i} &= \tau_2^{\sum_{i=I}^{k-1} 2i} = \tau_2^{2\left(\sum_{i=1}^{k-1} i - \sum_{i=1}^{I-1} i\right)} = \tau_2^{k(k-1)-I(I-1)} \\ &= \tau_2^{k^2-k-I^2+I} < \tau_2^{k^2-k-I^2} = \tau_2^{k(k-1)} \cdot \tau_2^{-I^2} = \tau_2^{k(k-1)} \cdot c_5\end{aligned}$$

thus

$$\begin{aligned}|\det \Delta_k|^2 &\leq c_3 c_2^{k-I} \prod_{i=I}^{k-1} \tau_2^{2i} < c_3 c_2^{k-I} \tau_2^{(k-1)^k} \cdot c_5 \\ &< c_3 c_6^k (\tau_2^{k-1})^k = c_3 (c_6 \tau_2^{k-1})^k = c_3 (c_6 \tau_2 \tau_2^k)^k = c_3 (c_4 \tau_2^k)^k.\end{aligned}$$

Because each $\det \Delta_k$ is an integer, they must be 0 for all k sufficiently large. Then Proposition 5.1 tells us that $f(z)$ is rational. □

5.2 Proof of Theorem 1.1

Lemma 5.3. *Let α be an algebraic number. Then* $\lceil \alpha^n \rceil = \lceil \alpha \rceil^n$ *for all integers* $n > 0$.

Proof. The Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts transitively on the conjugates of α, as well as on conjugates of α^n. By the multiplicativity of the automorphisms, the set of the conjugates of α^n equals the set of n-th powers of the conjugates of α. Hence $\lceil \alpha^n \rceil = \lceil \alpha \rceil^n$. □

Lemma 5.4. *Let α be a non-zero algebraic integer of degree d having minimal polynomial $P(z) = \prod_{i=1}^{d}(z - \alpha_i)$. Then the series*

$$\begin{aligned} F(1/z) &:= \sqrt{P_2^*\left(\frac{1}{z}\right)P_4^*\left(\frac{1}{z}\right)} \in \mathbb{C}[[1/z]] \\ &= \sqrt{\frac{1}{z^{2d}}P_2(z)P_4(z)} \\ &= \frac{1}{z^d}\sqrt{P_2(z)P_4(z)} \end{aligned}$$

lies in $\mathbb{Z}[[1/z]]$ and has zeros precisely at $\alpha_i^2, \alpha_i^4 (i = 1, \ldots, d)$, where the α_i are the conjugates of α.

Proof. Since the Maclaurin series of $(1+4Y)^{-1/2} = \sum_{j=0}^{\infty}(-1)\binom{2j}{j}Y^j \in \mathbb{Z}[Y]$, we also have $\sqrt{1+4Y} \in \mathbb{Z}[Y]$. Now from Lemma A.18 we have that $P_4(z) = P_2(z) + 4R(z)$ for some $R(z) \in \mathbb{Z}[z]$ of degree at most $d-1$. Then we obtain

$$\sqrt{P_2(z)P_4(z)} = \sqrt{P_2(z)(P_2(z) + 4R(z))} = P_2(z)\sqrt{1 + 4R(z)/P_2(z))}.$$

We now rewrite this identity in terms of reciprocal polynomials. So, on dividing by z^d , rewriting this identity in terms of the reciprocal polynomials P_2^*, P_4^* and R^* and using the fact that $P_2^*(0) = 1$ (because P_2 is monic we have that the constant term of P_2^* is equal to 1), we see that

$$\begin{aligned} F(1/z) &:= \sqrt{P_2^*\left(\frac{1}{z}\right)P_4^*\left(\frac{1}{z}\right)} \\ &= P_2^*\left(\frac{1}{z}\right)\sqrt{1 + 4z^{-(d-\deg R)}\frac{R^*\left(\frac{1}{z}\right)}{P_2^*\left(\frac{1}{z}\right)}} \in \mathbb{Z}\left[\left[\frac{1}{z^{d-\deg R}}\frac{R^*(1/z)}{P_2^*(1/z)}\right]\right] \end{aligned}$$

If we show

$$\frac{1}{z^{d-\deg R}}\frac{R^*(1/z)}{P_2^*(1/z)} \in \mathbb{Z}[[1/z]],$$

then $F(1/z) \in \mathbb{Z}[[1/z]]$. Now observe that

$$\frac{R^*(z)}{P_2^*(z)} = \frac{R^*(z)}{1-(1-P_2^*(z))} = \underbrace{R^*(z)\sum_{k\geq 0}(1-P_2^*(z))^k}_{\in\mathbb{Z}[[z]]}.$$

Thus

$$\frac{R^*(1/z)}{P_2^*(1/z)} \in \mathbb{Z}[[1/z]],$$

and

$$\frac{1}{z^{d-\deg R}}\frac{R^*(1/z)}{P_2^*(1/z)} \in \mathbb{Z}[[1/z]].$$

Therefore $F(1/z)$ lies in $\mathbb{Z}[[1/z]]$ and, because $P^*(1/z) = 1/z^d P(z)$, it has zeros precisely at $\alpha_i^2, \alpha_i^4 (i = 1, \ldots, d)$, where the α_i are the conjugates of α. ☐

Lemma 5.5. *Let $P(z) \in \mathbb{Z}[z]$ a monic integer irreducible polynomial and assume that P_2 is irreducible. If $\sqrt{P_2(z)P_4(z)}$ is a polynomial, then P is a cyclotomic polynomial.*

Proof. Since $P_2(z)$ is irreducible, for $\sqrt{P_2(z)P_4(z)}$ to be a polynomial, $P_4(z)$ must be equal to $P_2(z)$. Now take any root α of the irreducible polynomial P. The maximal modulus of a zero of $P_2(z)$ is $\overline{|\alpha|}^2$, while the maximal modulus of a zero of $P_4(z)$ is $\overline{|\alpha|}^4$. So $\overline{|\alpha|}^2 = \overline{|\alpha|}^4 = 1$ and, by Theorem A.19 α is a root of unity. ☐

Finally we can prove Theorem 1.1.

Proof of Theorem 1.1. Let $P(z) = \prod_{i=1}^d (z-\alpha_i)$ be the minimal polynomial of α and $F(1/z)$ be as in Lemma 5.4. First we construct a set containing all the non-regular points of $F(1/z)$. We take the $2d$ points $\alpha_1^2, \ldots, \alpha_d^2$ and $\alpha_1^4, \ldots, \alpha_d^4$ and we define the hedgehog $\mathcal{K} = \mathcal{K}(\alpha_1^2, \ldots \alpha_d^2, \alpha_1^4, \ldots, \alpha_d^4)$. Because these numbers have at most $2d'$ different arguments, at most $2d'$ of them are needed to define $\mathcal{K}$. So, by Corollary 4.5, $\mathcal{K}$ has transfinite diameter at most $e^{-0.39/2d'}\overline{|\alpha|}^4$.

Assume that $P_2(z)$ is irreducible and assume by contradiction $\overline{|\alpha|} < e^{0.39/8d'}$. Then $\mathcal{K}$ has transfinite diameter less than 1. Also we observe that $\alpha_1^2, \ldots, \alpha_d^2$ and $\alpha_1^4, \ldots, \alpha_d^4$ are the roots of $P_2(z)$ and $P_4(z)$ respectively, so the non-real roots come in pairs that are symmetric about the real axis. Thus $\mathcal{K}$ is symmetric about the real axis and contains the branch points of the square root function. Hence, $F(1/z) = \frac{1}{z^d}\sqrt{P_2(z)P_4(z)}$ is rational, by Theorem 5.2 , and having no poles except at $z = 0$, is in fact a polynomial in $1/z$. This implies that since $P_2(z)$ is irreducible and $z^d F(1/z) = \sqrt{P_2(z)P_4(z)}$ is a polynomial, by Lemma 5.5 α is a root of unity: a contradiction. We deduce that Theorem 1.1 is true whenever P_2 is irreducible.

Finally, for the case where P_2 is reducible assume that Theorem 1.1 is true for all non-zero algebraic integers of degree less than d, which are not roots of unity. The case when $d = 1$ is obvious. Now, since $P_2(z)$ is reducible, with α^2 one of its zeros, the degree of α^2 is less than that of α. So, by induction, the conclusion of the theorem holds for α^2. Also, since a pair $\alpha_j, -\alpha_j$ must occur among the conjugates of α, it follows from Galois theory that the arguments of the conjugates of α are in pairs $\theta, \theta + \pi$, say. Hence, the numbers α_j^2 have at most $\frac{1}{2}d'$ arguments between them. Thus, by induction $\overline{|\alpha^2|} \geq e^{0.39/8(d'/2)}$, or by Lemma 5.3 $\overline{|\alpha|} \geq e^{0.39/8d'}$. Thus the Theorem is true for α of degree d. □

A Appendix

A.1 Linear algebra

Theorem A.1 (Hadamard's Inequality)**.** *If $A = (a_{i,j})$ is a complex $n \times n$ matrix. Then*

$$|\det A| \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{i,j}|^2 \right)^{1/2}, \tag{A.1}$$

with equality if and only if either both sides are zero or $\sum_{i=1}^{n} a_{i,j}\overline{a_{ik}}$ for $j \neq k$.

Proof. Lemma 3.2 in [16]. □

Theorem A.2 (Cauchy-Binet Formula)**.** *Let A and B be matrices of size $n \times m$ and $m \times n$ respectively, with $n \leq m$. Then*

$$\det(AB) = \sum_{1 \leq k_1 < k_2 < \cdots < k_n \leq m} A_{k_1 \ldots k_n} B^{k_1 \ldots k_n},$$

where $A_{k_1 \ldots k_n}$ is the minor obtained from the columns of A whose numbers are $k_1, \ldots, k_n$ and $B^{k_1 \ldots k_n}$ is the minor obtained from the rows of B whose numbers are $k_1, \ldots, k_n$. In other words, det(AB) is the sum of products of the corresponding majors of A and B, where a major of a matrix is, by definition, a determinant of maximal order minor in the matrix.

Proof. We follow [13].
Let $C = AB, c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}$. Then

$$\begin{aligned}
\det C &= \sum_{\sigma \in S_n} (-1)^{\sigma} \left(\sum_{k_1} a_{1,k_1} b_{k_1,\sigma(1)} \cdots \sum_{k_n} a_{n,k_n} b_{k_n,\sigma(n)} \right) \\
&= \sum_{k_1,\ldots,k_n=1}^{m} a_{1,k_1} \cdots a_{n,k_n} \sum_{\sigma \in S_n} (-1)^{\sigma} b_{k_1,\sigma(1)} \cdots b_{k_n,\sigma(n)} \\
&= \sum_{k_1,\ldots,k_n=1}^{m} a_{1,k_1} \cdots a_{n,k_n} B^{k_1 \ldots k_n}.
\end{aligned}$$

The minor $B^{k_1 \ldots k_n}$ is non-zero only if the numbers $k_1, \ldots, k_n$ are distinct. Thus, the summation can be performed over distinct numbers $k_1 \ldots k_n$. Since

$B^{\tau(k_1)\dots\tau(k_n)} = (-1)^\tau B^{k_1\dots k_n}$ for any permutation τ of the numbers $k_1, \dots, k_n$, we have

$$\sum_{k_1,\dots,k_n=1}^{m} a_{1,k_1} \cdots a_{n,k_n} B^{k_1\dots k_n} = \sum_{k_1<k_2<\dots<k_n} \sum_{\tau} (-1)^\tau a_{1,\tau(1)} \cdots a_{n,\tau(n)} B^{k_1\dots k_n}$$
$$= \sum_{1\le k_1<k_2<\dots<k_n\le m} A_{k_1\dots k_n} B^{k_1\dots k_n}.$$

□

An alternative notation for the Cauchy-Binet formula is

$$\det(AB) = \sum_S A_S B^S$$

where the sum runs over the set of size n subsets of $\{1, \dots, m\}$ and A_S and B^S are defined as $A_{k_1\dots k_n}$ and $B^{k_1\dots k_n}$ for $\{k_1, \dots, k_n\} = S$.

Proposition A.3 (Vandermonde determinant)**.** *Let $X_1, X_2, \dots, X_n$ be indeterminates and let*

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_0 & X_1 & \cdots & X_n \\ X_0^2 & X_1^2 & \cdots & X_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ X_0^{n-1} & X_1^{n-1} & \cdots & X_n^{n-1} \end{pmatrix} \in Mat_{n+1,n+1}(\mathbb{Z}[X_0, \dots, X_n])$$

be the Vandermonde matrix. The determinant of this matrix is called the Vandermonde determinant and it is equal to

$$\det V = \prod_{1\le i<j\le n} (X_j - X_i).$$

Proof. If $X_k = X_l$ with $k \neq l$, then the Vandermonde determinant vanishes because in that case two rows of the determinant are identical. Hence, if $\det(V)$ is zero whenever $X_k = X_l$, $(X_k - X_l)$ divides the determinant as a polynomial in the X_i's. But that means that the product $\prod_{1\le i<j\le n}(X_j - X_i)$ must divide the determinant.

On the other hand, the determinant is a polynomial in the X_i's of degree at most $\binom{n}{2}$. Combined with the previous observation, this implies that the determinant equals the right-hand side product times, possibly, some constant. To compute the constant, we compare the coefficients of $X_1^0 X_2^1 \cdots X_n^{n-1}$ on both sides, this coefficient must be equal to one and this completes the proof. $\square$

Proposition A.4. *Let $X_1, X_2, \ldots, X_n$ be indeterminates. If $p_1, p_2, \ldots p_n$ are polynomials of the form $p_j(x) = a_j x^{j-1} +$ lower terms, then*

$$\det_{1\leq i,j\leq n}(p_j(X_i)) = a_1 a_2 \cdots a_n \prod_{1\leq i<j\leq n}(X_j - X_i).$$

This proposition can be proved in just the same way as the above proof of the Vandermonde determinant evaluation.

A.2 Schur polynomials

The following definitions and proposition are based on [19], [18] and [1].

Definition A.5. Fix integers $n \geq 1$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. The λ'th Schur function is the polynomial defined as

$$s_\lambda(x_1, x_2, \ldots, x_n) := \frac{|x_j^{\lambda_i+n-i}|_{1\leq i,j\leq n}}{|x_j^{n-i}|_{1\leq i,j\leq n}} = \frac{\begin{vmatrix} x_1^{\lambda_1+n-1} & x_2^{\lambda_1+n-1} & \ldots & x_n^{\lambda_1+n-1} \\ x_1^{\lambda_2+n-2} & x_2 x^{\lambda_2+n-2} & \ldots & x_n^{\lambda_2+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{\lambda_n} & x_2^{\lambda_n} & \ldots & x_n^{\lambda_n} \end{vmatrix}}{\begin{vmatrix} x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \ldots & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^0 & x_2^0 & \ldots & x_n^0 \end{vmatrix}}.$$

Definition A.6. Given a partition λ of n, a Young diagram of shape λ is an array of boxes arranged in rows. There are λ_i boxes in row i, each row of boxes starts at the leftmost position of the row and by the convention on partitions the lengths of rows are non-increasing.

Definition A.7. A Young tableau of shape λ is an assignment of the numbers $1, 2, \ldots, n$ to the n boxes of the Young diagram associated to λ.

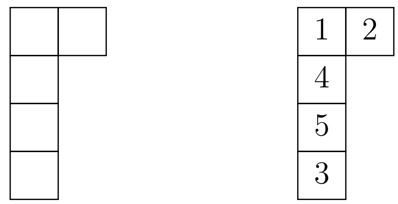

Figure 5: A Young diagram and Young tableau of shape $(2,1,1,1)$

Definition A.8. Let T be a tableau of shape λ. If the number of 1's, 2's , 3's ... occuring in T is $r_1, r_2, r_3 \ldots$ then associate the monomial $x^{w(T)} = x_1^{r_1} x_2^{r_2} x_3^{r_3} \cdots$ to T.

Proposition A.9. *$s_\lambda(x_1, x_2, \ldots, x_n) = \sum_T x^{w(T)}$ where the sum runs over all tableaux of shape λ.*

Proof. Corollary 12.5 in [18]. □

A.3 Bernoulli numbers and Bernoulli polynomials

Definition A.10. The Bernoulli numbers B_n are a sequence of signed rational numbers defined by the generating function

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n. \tag{A.2}$$

That is, we are to expand the left-hand side of this equation in powers of x, i.e., a Taylor series about x = 0. The coefficient of x^n in this expansion is $\frac{B_n}{n!}$.

Definition A.11. The Bernoulli polynomials $B_n(s)$ are defined by the generating function

$$F(x, s) = \frac{x e^{xs}}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n(s)}{n!} x^n.$$

We now prove a few properties of the Bernoulli numbers and polynomials.

Proposition A.12. *If $n \geq 3$ is odd, then $B_n = 0$.*

Proof. We extract the term for $n = 1$ from (A.2) and obtain

$$\frac{t}{e^t - 1} = -\frac{t}{2} + \sum_{\substack{n \geq 0 \\ n \neq 1}} B_n \frac{t^n}{n!}.$$

Thus

$$\frac{t}{e^t - 1} + \frac{t}{2} = \sum_{\substack{n \geq 0 \\ n \neq 1}} B_n \frac{t^n}{n!}.$$

We rewrite the left hand side as

$$\frac{t}{e^t - 1} + \frac{t}{2} = \frac{2t + t(e^t - 1)}{2(e^t - 1)} = \frac{t}{2} \cdot \frac{e^t + 1}{e^t - 1} = \frac{t}{2} \cdot \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}}$$

and observe that $\frac{t}{e^t-1} + \frac{t}{2}$ is an even function. Thus $\sum_{\substack{n \geq 0 \\ n \neq 1}} B_n \frac{t^n}{n!}$ is also even and therefore $B_n = (-1)^n B_n$ for $n \neq 1$. We conclude that if $n \geq 3$ and n is odd, then $B_n = 0$. □

Proposition A.13 (Faubauler's Formula)**.** *Let $i \geq 0$ be an integer and let $B_i = T^i + \cdots \in \mathbb{Q}[T]$ denote the i-th Bernoulli polynomial with constant term $B_i(0)$. Then*

$$\sum_{k=0}^{N-1} k^i = s_i(N) \text{ where } s_i = \frac{(B_{i+1}(T) - B_{i+1}(0))}{i+1} \in \mathbb{Q}[T].$$

A.4 Other results

Theorem A.14 (Newton's Identities)**.** *Take any complex variables $\alpha_1, \ldots, \alpha_d$ and define for $k \geq 0$*

$$s_k := \alpha_1^k + \cdots + \alpha_d^k,$$
$$t_k := (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq d} \alpha_{i_1} \cdots \alpha_{i_k}, \quad (\text{with } t_k = 0 \text{ if } k > d).$$

Then, for each $k \geq 1$, we have

$$s_k + \sum_{r=1}^{k-1} t_r s_{k-r} + k t_k = 0.$$

Proof. We have formally that

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{s_k}{z^k} &= \sum_{k=0}^{\infty} \frac{\alpha_1^k + \cdots + \alpha_d^k}{z^k} \\
&= \sum_{i=0}^{d} \sum_{k=0}^{\infty} \left(\frac{\alpha_i}{z} \right)^k \\
&= \sum_{i=1}^{d} \frac{1}{1 - \frac{\alpha_i}{z}} \\
&= z \sum_{i=1}^{d} \frac{1}{z - \alpha_i} \\
&= \frac{z \sum_{i=1}^{d} \prod_{\substack{j=1 \\ j \neq i}}^{n} (z - \alpha_j)}{\prod_{i=1}^{d} (z - \alpha_i)} \\
&= \frac{zP'(z)}{P(z)},
\end{aligned}$$

where

$$P(z) := \prod_{i=1}^{d} (z - \alpha_i) = z^d + t_1 z^{d-1} + t_2 z^{d-2} + \cdots + t_d.$$

Thus

$$\sum_{k=0}^{\infty} \frac{s_k}{z^k} P(z) = zP'(z),$$

$$\sum_{k=0}^{\infty} \frac{s_k}{z^k} (z^d + t_1 z^{d-1} + t_2 z^{d-2} + \cdots + t_d) = (d) z^d + (d-1) t_1 z^{d-1} + \cdots + t_{d-1} z.$$

Comparing the coefficients of z^{d-k} of $P(z) \sum_{k=0}^{\infty} (s_k / z^k)$ and $zP'(z)$, and using $s_0 = d$ finishes the proof. □

Theorem A.15. *Let* $P(z) = \prod_{i=1}^{d} (z - \alpha_i) \in \mathbb{Z}[z]$, *and*

$$s_k := \sum_{i=1}^{d} \alpha_i^k, \qquad (k = 1, 2, \dots)$$

as before. Then for all $n \in \mathbb{N}$ we have (with μ the Möbius function)

$$\sum_{l|n} s_l \mu\left(\frac{n}{l}\right) \equiv 0 \pmod{n}.$$

For the proof, we need the following simple lemma.

Lemma A.16. *Any integer power series*

$$1 + a_1 z + a_2 z^2 + \cdots + a_n z^n + \cdots$$

can be written formally as a product $\prod_{k=1}^{\infty}(1-z^k)^{b_k}$, where the exponents b_k are also all integers.

Proof. Inductively, if we have such a product for which the coefficients of $z^0, z^1, \ldots, z^{n-1}$ are the desired ones, say

$$\prod_{k=1}^{n-1}(1-z^k)^{b_k} \equiv 1 + a_1 z + a_2 z^2 + \cdots + a_{n-1} z^{n-1} + a'_n z^n \pmod{z^{n+1}}$$

say, then a'_n is an integer and if we multiply the product by $(1-z^n)^{a'_n - a_n}$ we obtain a product $\prod_{k=1}^{n}(1-z^k)^{b_k}$ for which the coefficients of $z^0, z^1, \ldots, z^{n-1}, z^n$ are the desired ones:

$$\begin{aligned}(1-z^n)^{a'_n - a_n}\prod_{k=1}^{n-1}(1-z^k)^{b_k} &\equiv 1 + a_1 z + \cdots + a_{n-1} z^{n-1} + a'_n z^n - a'_n z^n + a_n z^n \pmod{z^{n+1}} \\ &\equiv 1 + a_1 z + \cdots + a_{n-1} z^{n-1} + a_n z^n \pmod{z^{n+1}}.\end{aligned}$$

□

Proof of Theorem A.15. We use Lemma A.16 to write the reciprocal polynomial of $P(z)$ as

$$\prod_{i=1}^{d}(1-\alpha_i z) = \prod_{k=1}^{\infty}(1-z^k)^{b_k},$$

where the exponents b_k are all integers. Taking logarithms, we have

$$\sum_{i=1}^{d}\log(1-\alpha_i z) = \sum_{k=1}^{\infty} b_k \log(1-z^k).$$

Then, comparing coefficients of z_n in the Maclaurin expansions of both sides gives

$$-\frac{s_n}{n} = \sum_{k|n} -\frac{b_k}{n/k},$$

or $s_n = \sum_{k|n} kb_k$. Then, we apply Möbius inversion, Lemma A.17 to obtain $\sum_{l|n} s_l\mu\left(\frac{n}{l}\right) = nb_n$. ∎

Lemma A.17 (Möbius' Inversion Formula)**.** *Let f be an arithmetic function. Define $F(n) := \sum_{d|n} f(d)$ for $n \in \mathbb{Z}_{>0}$. Then*

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right)F(d) \quad \text{for } n \in \mathbb{Z}_{>0}.$$

Given a monic integer polynomial $P(z) = \prod_{i=1}^{d}(z - \alpha_i)$, let $P_l(z)$ denote the polynomial $\prod_{i=1}^{d}(z - \alpha_i^l)$. Because its coefficients are symmetric functions of the α_i, we have $P_l(z) \in \mathbb{Z}[z]$.

Lemma A.18. *For all $l, n \in \mathbb{Z}_{>0}$ we have*

$$\sum_{l|n} P_l(z)\mu\left(\frac{n}{l}\right) \equiv 0 \pmod{n}.$$

Proof. By Theorem A.15 the coefficient of z^{d-1} of $\sum_{l|n} P_l(z)\mu(n/l)$ is

$$-\sum_{l|n} s_l\mu(n/l) \equiv 0 \pmod{n}$$

Similarly, the coefficient of z^{d-j} of $\sum_{l|n} P_l(z)\mu(n/l)$ is also a multiple of n. This is seen by replacing P by the polynomial of degree $\binom{d}{j}$ whose zeroes are the products $\alpha_{i_1}\alpha_{i_2}\cdots\alpha_{i_j}$ for all possible j-element subsets $\{i_1, i_2, \ldots, i_j\}$ of $\{1, 2, \ldots, d\}$, and then considering its coefficient of $z^{\binom{d}{j}-1}$. ∎

Theorem A.19 (Kronecker's First Theorem)**.** *Suppose that α is a non-zero algebraic integer that lies, with its conjugates, in the unit disc $|z| \leq 1$. Then α is a root of unity.*

Proof. Suppose that α has degree d, with minimal polynomial

$$P_\alpha(z) = z^d + a_1 z^{d-1} + \cdots + a_d,$$

say. Then, because $(-1)^k a_k$ is the sum of all possible distinct k-tuples of the zeros of P, all of which have modulus at most 1, we have $|a_k| \leq \binom{d}{k}, (k = 0, \ldots, d)$. The same bounds apply to the coefficients of the polynomial P_r whose zeros are the r-th powers of those of P. Because the algebraic integers form a ring, α^r is also an algebraic integer, the coefficients of P_r are integers, so there are only finitely many possibilities for the coefficients of these polynomials. Hence, there are finitely many possibilities for all the zeros of all the P_r, and so in particular, there are only finitely many possibilities for the α^r. Therefore, two of them must be equal, say $\alpha^r = \alpha^s$ with $r < s$. Then $\alpha^{r-s} = 1$, and we see that α is a root of unity. $\square$

Lemma A.20. *Let $a_1, a_2, \ldots, a_n$ be real numbers and $b_1, b_2, \ldots, b_n$ positive real numbers. Then*

$$\frac{a_1^2}{b_1} + \cdots + \frac{a_n^2}{b_n} \geq \frac{(a_1 + \cdots + a_n)^2}{b_1 + \cdots + b_n}.$$

Proof. We first prove that for all real numbers a_1, a_2 and positive real numbers b_1, b_2

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \geq \frac{(a_1 + a_2)^2}{b_1 + b_2}. \tag{A.3}$$

Applying inequality (A.3) several times, we obtain

$$\begin{aligned} \frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} + \frac{a_3^2}{b_3} + \cdots + \frac{a_n^2}{b_n} &\geq \frac{(a_1 + a_2)^2}{b_1 + b_2} + \frac{a_3^2}{b_3} + \cdots + \frac{a_n^2}{b_n} \\ &\geq \frac{(a_1 + a_2 + a_3)^2}{b_1 + b_2 + b_3} + \cdots + \frac{a_n^2}{b_n} \\ &\geq \cdots \\ &\geq \frac{(a_1 + \cdots + a_n)^2}{b_1 + \cdots + b_n}. \end{aligned}$$

$\square$

Proposition A.21 (Stirling's approximation)**.** *For a positive integer n we have*

$$\log n! = n \log n - n + O(\log(n + 1)).$$

Proposition A.22 (Jensen's Inequality)**.** *For be a real concave function* ϕ*, real numbers* $x_1, x_2, \ldots, x_n$ *and positive real numbers* $a_1, \ldots, a_n$ *we have*

$$\frac{\sum_{i=1}^{n} a_i \phi(x_i)}{\sum_{i=1}^{n} a_i} \leq \phi\left(\frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i}\right).$$

References

[1] Jeremy Booher. Representations of the symmetric group via Young tableaux. 1931.

[2] Straus E. Cantor, D. On a conjecture of D. H. Lehmer. *Acta Arithmetica*, 42(1):97–100, 1982.

[3] Matt Davis. Transfinite diameter. *Ball State Undergraduate Mathematics Exchange*, 2(1):10–17, 2004.

[4] Vesselin Dimitrov. A proof of the Schinzel-Zassenhaus conjecture on polynomials. *arXiv preprint arXiv:1912.12545*, 2019.

[5] Edward Dobrowolski. Maximal modulus of conjugates of an algebraic integer. *BULLETIN DE L ACADEMIE POLONAISE DES SCIENCES-SERIE DES SCIENCES MATHEMATIQUES ASTRONOMIQUES ET PHYSIQUES*, 26(4):291–292, 1978.

[6] Edward Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arithmetica*, 34(4):391–401, 1979.

[7] Artūras Dubickas. On a conjecture of A. Schinzel and H. Zassenhaus. *Acta Arithmetica*, 63(1):15–20, 1993.

[8] Michael Fekete. Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen Koeffizienten. *Mathematische Zeitschrift*, 17(1):228–249, 1923.

[9] P HABEGGER. Separating roots of polynomials and the transfinite diameter. 2021.

[10] David Hilbert. Ein Beitrag zur theorie des legendreschen polynoms. In *Algebra· Invariantentheorie Geometrie*, pages 367–370. Springer, 1970.

[11] Christian Krattenthaler. Advanced determinant calculus. *The Andrews Festschrift*, pages 349–426, 1999.

[12] L. Kronecker. Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten. *Journal für die reine und angewandte Mathematik*, 53:173–175, 1857.

[13] Jin Ho Kwak and Sungpyo Hong. *Linear algebra.* Springer Science & Business Media, 2004.

[14] Derrick H Lehmer. Factorization of certain cyclotomic functions. *Annals of mathematics*, 34(3):461–479, 1933.

[15] Rudolf Lidl, Gary L Mullen, et al. Dickson polynomials. *(No Title)*, 1993.

[16] James McKee and Chris Smyth. Around the unit circle.

[17] Leonid Mirsky. *An introduction to linear algebra.* Courier Corporation, 2012.

[18] Amritanshu Prasad. An introduction to Schur polynomials. *arXiv preprint arXiv:1802.06073*, 2018.

[19] Robert A Proctor. Equivalence of the combinatorial and the classical definitions of Schur functions. *Journal of Combinatorial Theory-Series A*, 51(1):135–137, 1989.

[20] Andrzej Schinzel and Hans Zassenhaus. A refinement of two theorems of Kronecker. *Michigan Mathematical Journal*, 12:81–85, 1965.

[21] Chris J Smyth. On the product of the conjugates outside the unit circle of an algebraic integer. *Bulletin of the London Mathematical Society*, 3(2):169–175, 1971.