Algebra & Number Theory

Volume 18
2024
No. 11

msp

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2024 is US $525/year for the electronic version, and $770/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 18:11 (2024)
https://doi.org/10.2140/ant.2024.18.1945

Galois orbits of torsion points near atoral sets

Vesselin Dimitrov and Philipp Habegger

We prove that the Galois equidistribution of torsion points of the algebraic torus $\mathbb{G}_m^d$ extends to the singular test functions of the form $\log|P|$, where P is a Laurent polynomial having algebraic coefficients that vanishes on the unit real d-torus in a set whose Zariski closure in $\mathbb{G}_m^d$ has codimension at least 2. Our result includes a power-saving quantitative estimate of the decay rate of the equidistribution. It refines an ergodic theorem of Lind, Schmidt, and Verbitskiy, of which it also supplies a purely Diophantine proof. As an application, we confirm Ih's integrality finiteness conjecture on torsion points for a class of atoral divisors of $\mathbb{G}_m^d$.

1. Introduction

1A. ***Main results.*** Let $d \geq 1$ be an integer and let $\mathbb{G}_m^d$ denote the d-dimensional algebraic torus. We will identify $\mathbb{G}_m^d$ with $(\mathbb{C}\backslash\{0\})^d$, the group of its $\mathbb{C}$-points.

Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ be a *torsion point*, i.e., a point of a finite order. We define

$$\delta(\boldsymbol{\zeta}) = \inf\{|a| : a \in \mathbb{Z}^d\backslash\{0\} \text{ with } \boldsymbol{\zeta}^a = 1\} \tag{1-1}$$

where, here and throughout the article, $|\cdot|$ denotes the maximum-norm; we refer to Section 2 for the notation $\boldsymbol{\zeta}^a$.

MSC2010: primary 11J83; secondary 11R06, 14G40, 37A45, 37P30.
Keywords: number theory, Diophantine approximation, Mahler measure.

It is well-known that the Galois orbit

$$\{\boldsymbol{\zeta}^\sigma : \sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})\}$$

becomes equidistributed in $\mathbb{G}_m^d$ with respect to the Haar measure as $\delta(\boldsymbol{\zeta}) \to \infty$. More precisely, if $f : \mathbb{G}_m^d \to \mathbb{R}$ is a continuous function with compact support, then

$$\frac{1}{[\mathbb{Q}(\boldsymbol{\zeta}) : \mathbb{Q}]} \sum_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})} f(\boldsymbol{\zeta}^\sigma) \to \int_{[0,1)^d} f(\boldsymbol{e}(x))\, dx \tag{1-2}$$

as $\delta(\boldsymbol{\zeta}) \to \infty$ where

$$\boldsymbol{e}(x) = (e^{2\pi\sqrt{-1}x_1}, \dots, e^{2\pi\sqrt{-1}x_d}) \tag{1-3}$$

for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$.

Our aim is to investigate the equidistribution result for test functions $f = \log|P|$ where P is a nonzero Laurent polynomial in d unknowns and with algebraic coefficients. Such P may vanish on $(S^1)^d$, where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle, and so f is not defined everywhere. But for $\delta(\boldsymbol{\zeta})$ large in terms of P, Laurent's theorem [1984] also known as the Manin–Mumford conjecture for $\mathbb{G}_m^d$, implies that P does not vanish at any conjugate of $\boldsymbol{\zeta}$; see also [Sarnak and Adams 1994] for another proof. Moreover, the integral of f over $(S^1)^d$ exists as the singularity is merely logarithmic. It is known as the *Mahler measure*

$$m(P) = \int_{[0,1)^d} \log|P(\boldsymbol{e}(x))|\, dx, \tag{1-4}$$

see for instance Section 3.4 in [Schinzel 2000] for the convergence of this integral for arbitrary $P \in \mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$.

A *torsion coset* of $\mathbb{G}_m^d$ is the translate of a connected algebraic subgroup of $\mathbb{G}_m^d$ by a point of finite order. We call a torsion coset *proper* if it does not equal $\mathbb{G}_m^d$.

We call $P \in \mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$ *essentially atoral* if the Zariski closure of

$$\{(z_1, \dots, z_d) \in (S^1)^d : P(z_1, \dots, z_d) = 0\}$$

in $\mathbb{G}_m^d$ is a finite union of irreducible algebraic sets of codimension at least 2 and proper torsion cosets.

For example, if $d = 1$ then P is essentially atoral if and only if it does not vanish at any point of infinite multiplicative order in S^1.

Lind, Schmidt and Verbitskiy [Lind et al. 2013, Definition 2.1] define the notion of an *atoral* Laurent polynomial $P \in \mathbb{Z}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$. An atoral Laurent polynomial is essentially atoral in our sense. Moreover, if P is irreducible then it is atoral if and only if the intersection of its zero locus with $(S^1)^d$ has dimension at most $d - 2$ as a semialgebraic set, see Proposition 2.2 [Lind et al. 2013]. A related, but not quite equivalent, definition of atoral Laurent polynomials with complex coefficients was introduced earlier by Agler, McCarthy and Stankus [Agler et al. 2006].

Let $\overline{\mathbb{Q}}$ denote the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. We are ready to state our first result.

Theorem 1.1. *For each essentially atoral $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$ there exists $\kappa > 0$ with the following property. Suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order with $\delta(\boldsymbol{\zeta})$ sufficiently large. Then $P(\boldsymbol{\zeta}^\sigma) \neq 0$ for all $\sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})$ and*

$$\frac{1}{[\mathbb{Q}(\boldsymbol{\zeta}) : \mathbb{Q}]} \sum_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})} \log|P(\boldsymbol{\zeta}^\sigma)| = m(P) + O(\delta(\boldsymbol{\zeta})^{-\kappa})$$

as $\delta(\boldsymbol{\zeta}) \to \infty$, where the implicit constant depends only on d and P.

Theorem 8.8 below is a more precise version of this result. In particular, we allow σ to range over subgroups of $\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})$ whose index and conductor grow sufficiently slow, the conductor is defined in Section 3. Moreover, κ depends only on d and the number of nonzero terms appearing in P. Our method of proof allows one to determine an explicit value for κ.

Torsion points in $\mathbb{G}_m^d$ are characterized as the algebraic points of height zero; see Section 2 for the definition of the height $h : \mathbb{G}_m^n(\overline{\mathbb{Q}}) \to [0, \infty)$. Bilu [1997] proved that Galois orbits of algebraic points $\boldsymbol{\alpha} \in \mathbb{G}_m^d$ of small height satisfy an analogous equidistribution statement as (1-2), asymptotically as $h(\boldsymbol{\alpha}) \to 0$ and $\delta(\boldsymbol{\alpha}) \to \infty$; the definition (1-1) extends naturally to nontorsion points and may take infinity as a value. It is natural to ask whether Theorem 1.1 admits a suitable generalization to points of small height. Autissier's example [2006] rules out the verbatim generalization already for $\mathbb{G}_m$. He constructed a sequence $(\alpha_n)_{n\in\mathbb{N}}$ of pairwise distinct algebraic numbers whose height tends to 0 but such that $(1/[\mathbb{Q}(\alpha_n) : \mathbb{Q}]) \sum_\sigma \log|\sigma(\alpha_n) - 2|$ tends to 0 for $n \to \infty$. But the integral of the corresponding test function against the unit circle is $\log 2$. An interesting problem still arises if the test function has at worst a logarithmic singularity of real codimension at least 2 on $(S^1)^d$. Suppose that $|f(z)|$ is $O\big(\big|\log(|P(z)|^2 + |Q(z)|^2)\big|\big)$ on an open neighborhood of $(S^1)^d$ in $\mathbb{G}_m^d$, where P and Q are nonconstant and coprime Laurent polynomials with algebraic coefficients, and that f vanishes on the complement of a compact set in $\mathbb{G}_m^d$. One may then ask about comparing the average of f over the Galois orbit of $\boldsymbol{\alpha} \in \mathbb{G}_m^d(\overline{\mathbb{Q}})$ with the average of f over $(S^1)^d$: is their difference bounded by $\ll_f (h(\boldsymbol{\alpha}) + \delta(\boldsymbol{\alpha})^{-1})^\kappa$, for some $\kappa > 0$ depending only on P and Q? We also mention Chambert-Loir and Thuillier's Théorème 1.2 [2009] which is a general equidistribution result for points of small height, allowing $\log|P|$ as a test function if the zero locus of P in $\mathbb{G}_m^d$ is a finite union of torsion cosets. In this paper we allow $\log|P|$ as a test function if P is essentially atoral but we average over points of finite order.

Our Theorem 1.1 recovers a variant of the result of Lind, Schmidt and Verbitskiy [2013]. In their work, the sum is not over the Galois orbit of a single point of finite order but rather over a finite subgroup G of $\mathbb{G}_m^d$. For this purpose we define

$$\delta(G) = \inf\{|a| : a \in \mathbb{Z}^d\backslash\{0\} \text{ such that } \boldsymbol{\zeta}^a = 1 \text{ for all } \boldsymbol{\zeta} \in G\}. \tag{1-5}$$

Each finite subgroup of $\mathbb{G}_m^d$ is a disjoint union of Galois orbits. This observation allows us to recover the theorem of Lind, Schmidt, and Verbitskiy with an estimate on the decay rate.

Theorem 1.2. *Let $P \in \mathbb{Q}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$ be essentially atoral. There exists $\kappa > 0$ such that for any finite subgroup $G \subset \mathbb{G}_m^d$ we have*

$$\frac{1}{\#G} \sum_{\substack{\boldsymbol{\zeta} \in G \\ P(\boldsymbol{\zeta}) \neq 0}} \log|P(\boldsymbol{\zeta})| = m(P) + O(\delta(G)^{-\kappa}) \tag{1-6}$$

where the implicit constant depends only on d and P.

To relate (1-6) to the expression in Lind, Schmidt, and Verbitskiy's Theorem 1.3 [Lind et al. 2013] we refer to [Lind et al. 2010, Lemma 2.1] as well as the comments on pages 1063 and 1064 of [Lind et al. 2013]. Note that G is Ω_Γ and $\#G$ is $|\mathbb{Z}^d/\Gamma|$ in the notation of [Lind et al. 2013].

Lind, Schmidt, and Verbitskiy's approach is based on an in-depth study [Schmidt and Verbitskiy 2009; Lind et al. 2010; 2013] of an associated dynamical system: the algebraic $\mathbb{Z}^d$-action on a closed, shift-invariant subgroup of $(S^1)^{\mathbb{Z}^d}$ whose dual is $\mathbb{Z}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]/(P)$. The atoral condition, in the sense of [Lind et al. 2013], turns out to be equivalent to the existence of a nontrivial summable homoclinic point.

Theorem 1.2 may be read as a strong quantitative estimate on the growth of periodic points for such dynamical systems. The refinement to Galois orbits, Theorem 1.1, does not seem to be directly possible by the homoclinic method, nor does it seem to follow formally from the case (1-6) of finite subgroups, which is where the dynamical method applies.

Our method of proof draws its origins in work of Duke [2007]. It differs from the method of Lind, Schmidt, and Verbitskiy. However, it is striking that the notion of atoral appears crucially in both approaches.

The first-named author [Dimitrov 2016] was able to prove Theorem 1.2 for a general Laurent polynomial when G equals the group of N-torsion elements in $\mathbb{G}_m^d$.

Let us return to Galois orbits. We believe that the hypothesis on P being essentially atoral is also unnecessary in Theorem 1.1 on Galois orbits. The next conjecture sums up our expectations. It is related to Schmidt's conjecture [1995, Remark 21.16(2)].

Conjecture 1.3. *For each $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$ there exists $\kappa > 0$ with the following property. Suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order with $\delta(\boldsymbol{\zeta})$ sufficiently large. Then $P(\boldsymbol{\zeta}^\sigma) \neq 0$ for all $\sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})$ and*

$$\frac{1}{[\mathbb{Q}(\boldsymbol{\zeta}) : \mathbb{Q}]} \sum_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/\mathbb{Q})} \log|P(\boldsymbol{\zeta}^\sigma)| = m(P) + O(\delta(\boldsymbol{\zeta})^{-\kappa})$$

as $\delta(\boldsymbol{\zeta}) \to \infty$, where the implicit constant depends only on d and P.

For $d = 1$ this conjecture follows from work of M. Baker, Ih, and Rumely [Baker et al. 2008], see their statement around (6). They use a version of A. Baker's deep estimates on linear forms in logarithms. Our Theorem 1.1 in the case $d = 1$ does not cover polynomials that vanish at a point of infinite order on the unit circle and therefore avoids the use of linear forms in logarithms. The conjecture is open already for $d = 2$ and

$$P = X_1 + X_1^{-1} + X_2 + X_2^{-1} - 3.$$

1B. ***Ih's conjecture on integral torsion points.*** As another application of our results we derive a special case of Ih's conjecture [Baker et al. 2008] in the multiplicative setting. Let $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$. Ih's conjecture predicts that the set of torsion points $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ such that $P(\boldsymbol{\zeta})$ is an algebraic unit is not Zariski dense in $\mathbb{G}_m^d$, unless the zero set of P in $\mathbb{G}_m^d$ is itself a finite union of proper torsion cosets. M. Baker, Ih, and Rumely [2008] cover the case $d = 1$ for arbitrary polynomials using their work on Conjecture 1.3. Here we mimic the approach of M. Baker, Ih, and Rumely and solve a case of Ih's conjecture for essentially atoral polynomials with integral coefficients and any d.

Corollary 1.4. *Let $K \subset \mathbb{C}$ be a number field with ring of integers $\mathbb{Z}_K$ and let $P \in \mathbb{Z}_K[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$. Suppose that the zero set of P in $\mathbb{G}_m^d$ is not a finite union of torsion cosets. Suppose in addition that $\tau(P)$ is essentially atoral for all field embeddings $\tau : K \to \mathbb{C}$. Then there exists $B \geq 1$ such that if $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order and $P(\boldsymbol{\zeta})$ is an algebraic unit, then $\delta(\boldsymbol{\zeta}) \leq B$.*

Ih's conjecture expects the existence of B without assuming that each $\tau(P)$ is essentially atoral. Observe that the result of M. Baker, Ih, and Rumely is not a direct consequence of this corollary, as we do not allow univariate polynomials that vanish at a point of infinite multiplicative order on the unit circle. Our approach does not depend on the theory of linear forms in logarithms.

A special class of atoral polynomials, to which our results apply *a fortiori*, are the irreducible integer Laurent polynomials $P \in \mathbb{Z}[X_1^{\pm 1}, \dots, X_n^{\pm 1}] \setminus \{0\}$ that are not fixed up to a monomial factor and up to a sign by the involution sending each X_i to $1/X_i$. We call these P asymmetric. They are atoral in the sense of Lind, Schmidt and Verbitskiy, see the proof of [Lind et al. 2013, Proposition 2.2]. Hence an asymmetric Laurent polynomial is essentially atoral. The converse is false as the Laurent polynomial

$$X_1 + X_1^{-1} + X_2 + X_2^{-1} - 4.$$

is essentially atoral; indeed, its zero locus on $(S^1)^2$ consists of the single point $(1, 1)$.

If $K = \mathbb{Q}$, Corollary 1.4 in the case of an asymmetric, and thus necessarily irreducible Laurent polynomial P, can be deduced as follows from the Manin–Mumford conjecture for $\mathbb{G}_m^d$. Indeed, if γ is a unit in the ring of algebraic integers of a cyclotomic field, then $\eta = \bar{\gamma}/\gamma$ is an algebraic integer whose Galois conjugates lie on S^1. So η is a root of unity by Kronecker's theorem, see [Bombieri and Gubler 2006, Theorem 1.5.9]. We consider the zero $(\eta, \boldsymbol{\zeta})$ of $P(X_1^{-1}, \dots, X_d^{-1}) - X_0 P(X_1, \dots, X_d)$, which is irreducible and defines an algebraic subset of $\mathbb{G}_m^{d+1}$ none of whose geometric irreducible components is a torsion coset. A similar argument applies if K is a totally real number field.

1C. ***Overview of the proof.*** We close the introduction by describing the method of proof of Theorem 1.1, which builds upon work of the second-named author [Habegger 2018] and is related to the approach of Duke [2007]. The basic idea is to reduce the multivariate statement in Theorem 1.1 to the univariate case. Whereas we worked with torsion points of prime order in [Habegger 2018], a new technical difficulty in this paper is that we allow torsion points of arbitrary order.

Any torsion point $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ of order N takes on the form $(\zeta^{a_1}, \dots, \zeta^{a_d})$ where $\zeta = \boldsymbol{e}(1/N)$ is a root of unity of order N and $a = (a_1, \dots, a_d) \in \mathbb{Z}^d$. The precise manner how the nonunique a is chosen is

delicate and will be discussed below. The notation $\boldsymbol{\zeta} = \zeta^a$ will be quite useful. A nonboldface ζ denotes a root of unity and boldface $\boldsymbol{\zeta}$ suggests a torsion point of $\mathbb{G}_m^d$.

If P is as in Theorem 1.1, but for simplicity with coefficients in $K = \mathbb{Q}$, we define the univariate polynomial

$$Q(X) = P(X^a) = P(X^{a_1}, \dots, X^{a_d}) \in \mathbb{Q}[X^{\pm 1}]. \tag{1-7}$$

Multiplying Q by a power of X turns out to be harmless, so one can assume that Q is a polynomial. The values $|P(\boldsymbol{\zeta}^\sigma)|$ equal the values of $|Q(\zeta^\sigma)|$ as σ ranges over $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$.

The univariate case and root separation (Section 4)**.** Let us suppose for the moment that $\boldsymbol{\zeta} = \zeta$ is a root of unity. It is classical that the Galois conjugates of ζ are equidistributed around the unit circle; we recall of these facts in Section 3. So (1-2) holds for $f(z) = \log|Q(z)|$ provided Q has no zero on the unit circle. In Proposition 4.5 we make convergence quantitative for such Q. Roughly speaking, for all $\epsilon > 0$ we have

$$\frac{1}{[\mathbb{Q}(\zeta) : \mathbb{Q}]} \sum_\sigma \log|Q(\zeta^\sigma)| = m(Q) + O_{P,\epsilon}\left(\frac{|a|^{1+\epsilon}}{N^{1-\epsilon}}\right) \tag{1-8}$$

where σ runs over $\mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$. Actually, the hypothesis on Q is slightly weaker as we allow it to vanish at roots of unity, if all $Q(\zeta^\sigma) \neq 0$. This hypothesis is ultimately a reflection of the hypothesis that the multivariate P is essentially atoral in Theorem 1.1. Indeed, in the univariate case, being essentially atoral boils down to not vanishing at any point of infinite multiplicative order in S^1. The hypothesis on Q is crucial for our method to work. The main difficulty we encounter in the average (1-8) are exceptionally small values of Q at some ζ^σ. The burden is to show, in a uniform sense, that no complex root z of Q can be too close to ζ^σ in a suitable sense.

If z is itself a root of unity, doing this is straightforward as $|z - 1| \gg 1/\operatorname{ord}(z)$.

The difficulty lies in the case when z has infinite multiplicative order. Here it is tempting to apply a version of Baker's theorem on linear forms in logarithms, as did M. Baker, Ih, and Rumely [2008]. However, and as already discussed by Duke [2007, Section 3] this seems unhelpful for the problem at hand. Indeed, estimates on linear forms in two logarithms such as [Laurent et al. 1995] lead to a factor $[\mathbb{Q}(z) : \mathbb{Q}]^2 = O(|a|^2)$ in a bound for any member of the sum in (1-8). This is not good enough for our application as $|a|^2/[\mathbb{Q}(\zeta) : \mathbb{Q}]$ may spoil the average in (1-8).

Our solution is to use the banal inequality $|z - \zeta| \geq \big||z| - 1\big|$ which lies at the heart of the method here and in [Habegger 2018]. As z is no root of unity, and as Q does not vanish at points of infinite multiplicative order on S^1, we have $|z| \neq 1$ and so the banal inequality provides a nontrivial lower bound. We now explain how it leads to a useful estimate on $|z - \zeta|$ via lower bounding $\big||z| - 1\big|$.

If z is close to the unit circle, then $\big||z| - 1\big|$ is approximately $|z - 1/\bar{z}|$. In [Habegger 2018] a result of Mahler [1964] on the separation of roots of an integer polynomial led to a suitable lower bound for $|z - 1/\bar{z}|$. In that paper, Habegger used his counting result on approximations to a set definable in an o-minimal structure. This allowed to make Mahler's estimate uniform over the various zeros z of Q.

The main tool of the present paper is a uniform generalization of Mahler's inequality for the separation of several pairs of roots of Q. Such a generalization was obtained by Mignotte [1995]. In Section 4 we give a variant of Mignotte's theorem that is tailored to our application and is self-contained. We thus bypass the o-minimal theory used in [Habegger 2018]. We still require Bombieri, Masser, and Zannier's theorem [Bombieri et al. 2007] to be mentioned below. Moreover, our Theorem 1.1 is effective in nature.

A possible approach towards Conjecture 1.3 lies in extending (1-8) to Q that are allowed to vanish at any point of S^1. As observed, we lack a suitable lower bound for $|z-\zeta|$ if z is an algebraic number of infinite multiplicative order on the unit circle. As suggested in the similar setting of [Habegger 2018, Lemma 4.2], it turns out that the z of interest have small height $h(z)$. We therefore propose the following conjecture.

Conjecture 1.5. *For all $B \geq 1$ and $\epsilon > 0$ there exists a constant $c = c(B,\epsilon) > 0$ with the following property. Let $z \in \mathbb{C}$ be an algebraic number with $|z| = 1$ and $h(z) \leq B/D$ where $D = [\mathbb{Q}(z) : \mathbb{Q}]$. If $\zeta \in \mathbb{C}\backslash\{z\}$ is a root of unity of order N, then $\log|\zeta - z| \geq -cD^{1+\epsilon}N^{\epsilon}$.*

The crux of this conjecture is its dependency on the degree D. In comparison, the state-of-the-art results in the theory of linear forms in two logarithms of algebraic numbers in the D-aspect, such as Laurent, Mignotte, and Nesterenko's Théorème 3 [Laurent et al. 1995], have only a quadratic dependency on D.

Equidistribution of torsion points **(Section 3).** We return to the case $\boldsymbol{\zeta} = \zeta^a$ of a general torsion point in $\mathbb{G}_m^d$ of order N. The exponent vector a used to define Q as in (1-7) depends on $\boldsymbol{\zeta}$. For this reason it is important that the error term in Proposition 4.5 is explicit in terms of Q. Moreover, it is important to choose a with $|a|$ as small as possible. For fixed ζ the exponent a is well-defined up to addition of an element in $N\mathbb{Z}^d$. So clearly we may assume $|a| \leq N$, although this is not good enough in view of (1-8). Fortunately, there is a second degree of freedom, namely we can replace $\boldsymbol{\zeta}$ by any Galois conjugate of itself.

This leads us to classical questions of equidistribution of the Galois orbit of $\boldsymbol{\zeta}$; we compile the necessary statements in Section 3. Using the Erdös–Turán theorem and the theory of Gauss sums, Lemma 3.7 produces a with $|a| = O(N\delta(\boldsymbol{\zeta})^{-1/(3d)})$ such that ζ^a is a Galois conjugate of $\boldsymbol{\zeta}$.

Let us return to the error term in (1-8). One factor N cancels out and the error term becomes $N^{2\epsilon}\delta(\boldsymbol{\zeta})^{-(1+\epsilon)/(3d)}$. The innocuous ϵ in (1-8) is ultimately responsible for the factor $N^{2\epsilon}$. Although $\delta(\boldsymbol{\zeta}) \leq N$, there is no nontrivial bound in the reverse direction and $N^{2\epsilon}\delta(\boldsymbol{\zeta})^{-(1+\epsilon)/(3d)}$ could explode.

Factoring $\boldsymbol{\zeta}$ **(Section 5).** The solution to this problem is described in Section 5. In Proposition 5.1 we factor $\boldsymbol{\zeta}$ into a product $\boldsymbol{\eta}\boldsymbol{\xi}$ where $\boldsymbol{\xi}$ has finite order M such that $\boldsymbol{\xi} = \boldsymbol{e}(a/M)$ where $|a| = O(M^{1-\kappa})$. Moreover, the order of $\boldsymbol{\eta}$ is bounded from above by a small power of N. The power saving obtained in the exponent of N is small even when compared to the saving obtained for $|a|$. The methods employed come from the geometry of numbers and slopes of lattices in $\mathbb{R}^d$.

We will replace $\boldsymbol{\zeta}$ by $\boldsymbol{\xi}$ and the univariate polynomial $Q(X) = P(X^a)$ by $P(\boldsymbol{\eta}^a X^a)$. This last transformation does not change the height or the monomial structure of Q. But it can change the field generated by its coefficients as the order of $\boldsymbol{\eta}$ and hence its field of definition vary as $\boldsymbol{\zeta}$ varies. For this reason, we must keep track of the base field of Q throughout the whole argument.

Putting everything together (Sections 6, 7, 8)**.** In Sections 6 and 8 we put all ingredients together to prove the final result. Here we apply a result of Bombieri, Masser, and Zannier [2007] on the intersections of a subvariety in $\mathbb{G}_m^d$ of codimension at least 2 with all 1-dimensional algebraic subgroups of $\mathbb{G}_m^d$. Roughly speaking, this result shows that if P is essentially atoral, then for "most" choices of a the univariate polynomial Q as in (1-7) does not vanish at any point of infinite multiplicative order on S^1. Recall that this property of Q was crucial to deduce (1-8). Bombieri, Masser, and Zannier's result is related to the study of unlikely intersections, for an overview we refer to Zannier's book [2012]. Another tool that makes an appearance is Lawton's theorem [1983].

The intermediate Section 7 contains a weak version of a result of Hlawka [1971] on the numerical integration of a continuous, multivariate function. The results obtained there are useful in connection with the function attaching the Mahler measure to a nonzero polynomial.

Appendices. In Appendix A we give a quantitative version of Lawton's theorem [1983] regarding the convergence of a sequence of Mahler measures. Unfortunately, we are not able to use the very closely related theorem in [Habegger 2018] as we require additional uniformity. The arguments in this appendix follow closely Lawton's strategy. After this paper was submitted, Brunault, Guilloux, Mehrabdollahi, and Pengo proved a higher dimensional generalization of our version of Lawton's theorem with explicit constants [Brunault et al. 2022].

Finally, in Appendix B we show how to deduce Theorem 1.2, the theorem of Lind, Schmidt and Verbitskiy, from our Theorem 1.1.

1D. ***Final remarks.*** The results mentioned above, in particular the theorem of Bombieri, Masser, and Zannier, also play an important role in Le's approach [2014]. The question on how small a sum of roots of unity can be was raised by Myerson [1986] in connection with a combinatorial question [Myerson 1979; 1980] which was later studied by Duke [2007]. Dubickas [2018] has more recent work in this direction for sums of 2 and 3 roots of unity of prime order.

2. Notation and preliminaries

Apart from the notation already introduced we use $\mathbb{N}$ to denote the natural numbers $\{1, 2, 3, \dots\}$. If $x = (x_1, \dots, x_m)$ with all x_i elements in an abelian group G and if $A = (a_{i,j})_{i,j} \in \mathrm{Mat}_{m,n}(\mathbb{Z})$ we write $x^A = (x_1^{a_{1,1}} \cdots x_m^{a_{m,1}}, \dots, x_1^{a_{1,n}} \cdots x_m^{a_{m,n}}) \in G^n$. So if $B \in \mathrm{Mat}_{n,p}(\mathbb{Z})$, then $(x^A)^B = x^{AB}$. For a commutative ring R with 1 we let $R^\times$ denote its group of units. Euler's function φ maps $N \in \mathbb{N}$ to the cardinality of $(\mathbb{Z}/N\mathbb{Z})^\times$. The group of all roots of unity in $\mathbb{C}^\times$ is μ_∞. We often identify $\mathbb{G}_m^d$ with the set of its complex points $(\mathbb{C}^\times)^d$ and let 1 denote the unit element $(1, \dots, 1) \in \mathbb{G}_m^d$. If $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ is a torsion point, we write

$\operatorname{ord}(\boldsymbol{\zeta})$ for its order. We write $\langle \cdot, \cdot \rangle$ for the Euclidean inner product on $\mathbb{R}^d$, $|\cdot|_2$ for the Euclidean norm on $\mathbb{R}^d$, and $|\cdot|$ for the maximum-norm on $\mathbb{R}^d$ and $\mathrm{Mat}_{m,n}(\mathbb{R})$. We define $\log^+ x = \log \max\{1, x\}$ for all $x \geq 0$.

The constants implicit in Vinogradov's notation $\ll_{x,y,z,\ldots}$, $\gg_{x,y,z,\ldots}$, and in $O_{x,y,z,\ldots}(\cdots)$ depend only on the values $x, y, z, \ldots$ appearing in the subscript.

Let $P \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_d^{\pm 1}]\backslash\{0\}$, then $|P|$ denotes the maximum-norm of the coefficient vector of P and we set $|0| = 0$. Recall that $m(P)$ is the Mahler measure of P. It follows from Corollaries 4 and 6 in Chapter 3.4 of [Schinzel 2000] that $\exp(m(P))$ is at most the Hermitian norm of the coefficient vector of P. Suppose P has at most $k \geq 1$ nonzero terms, we find

$$m(P) \leq \log|P| + \tfrac{1}{2}\log k. \tag{2-1}$$

The following result of Dobrowolski and Smyth [2017, Corollary 2] provides a reverse inequality of the same quality.

Theorem 2.1 (Dobrowolski and Smyth). *Suppose $P \in \mathbb{C}[X_1^{\pm 1}, \ldots, X_d^{\pm 1}]\backslash\{0\}$ has at most $k \geq 2$ nonzero terms with k an integer. Then $m(P) \geq \log|P| - (k-2)\log 2$.*

Therefore,

$$|m(P) - \log|P|| \ll k \tag{2-2}$$

with absolute implied constant. Observe that if P is a polynomial, then

$$m(P) \geq \log|P| - \log(2)\sum_{i=1}^{d} \deg_{X_i} P$$

by the classical Lemma 1.6.10 of [Bombieri and Gubler 2006]. So (2-2) is stronger when the number of terms in P is known to be bounded, which is often the case in our work.

Let x be an element of a number field K. The absolute logarithmic Weil height, or just height, of x is

$$h(x) = \frac{1}{[K:\mathbb{Q}]}\sum_v [K_v : \mathbb{Q}_v] \log \max\{1, |x|_v\}; \tag{2-3}$$

here v runs over all places of K normalized such that $|2|_v = 2$ for an infinite place v and $|p|_v = 1/p$ if v lies above the rational prime p, the completion of K with respect to v is K_v and the completion of $\mathbb{Q}$ with respect to the restriction of v is $\mathbb{Q}_v$. Let P be a nonzero Laurent polynomial with coefficients $x_0, \ldots, x_n \in K$. The absolute logarithmic Weil height, or just height, of P is

$$h(P) = \frac{1}{[K:\mathbb{Q}]}\sum_v [K_v : \mathbb{Q}_v] \log \max\{|x_0|_v, \ldots, |x_n|_v\}. \tag{2-4}$$

See [Bombieri and Gubler 2006, Chapter 1] for more details on heights. For example, $h(x)$ and $h(P)$ are well-defined for $x \in \overline{\mathbb{Q}}$ and $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \ldots, X_d^{\pm 1}]$, i.e., the values do not depend on the number field K containing x and the coefficients of P, respectively. Moreover $h(P) = h(\lambda P)$ for all $\lambda \in \overline{\mathbb{Q}}^\times$.

3. Quantitative Galois equidistribution for torsion points

We need a strong enough quantitative version of the Galois equidistribution of torsion points $\boldsymbol{\zeta}$ of $\mathbb{G}_m^d$, with a power saving discrepancy in $\delta(\boldsymbol{\zeta})$ defined in (1-1).

Different approaches are possible and we opt to use the Erdös–Turán–Koksma bound. This reduces the problem to the estimation of certain exponential sums, which happen to be Gauss sums that can be explicitly evaluated.

Let $N \in \mathbb{N}$. For a divisor $f \in \mathbb{N}$ of N we work with the canonical surjective, homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/f\mathbb{Z})^\times$ induced by reducing modulo f.

The *conductor* f_G of a subgroup G of $(\mathbb{Z}/N\mathbb{Z})^\times$ is the least positive integer $f \mid N$ such that G contains $\ker((\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/f\mathbb{Z})^\times)$.

Certainly, f_G is well-defined as $\ker((\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N\mathbb{Z})^\times)$ is the trivial subgroup. Moreover, $f_{(\mathbb{Z}/N\mathbb{Z})^\times} = 1$. But one should take care as the conductor of $G = \{1\}$ is $N/2$ for $N \equiv 2 \pmod 4$. Observe that $[(\mathbb{Z}/N\mathbb{Z})^\times : G] \le \varphi(f_G)$.

The group $(\mathbb{Z}/N\mathbb{Z})^\times$ is naturally isomorphic to the Galois group of $\mathbb{Q}(\zeta)/\mathbb{Q}$, where ζ is a root of unity of order N. Let $L \subset \mathbb{Q}(\zeta)$ be the fixed field of G. Then L lies in the fixed field of $\ker((\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/f_G\mathbb{Z})^\times)$ which equals $\mathbb{Q}(\zeta_{f_G})$ where ζ_{f_G} is a root of unity of order f_G.

Let $f \ge 1$ be an integer and ζ_f of order f. We claim $L \subset \mathbb{Q}(\zeta_f)$ if and only if $f_G \mid f$. Indeed, if the inclusion holds, then $L \subset \mathbb{Q}(\zeta_f) \cap \mathbb{Q}(\zeta_{f_G})$. It is well-known that the intersection is generated by a root of unity of order $\gcd(f, f_G)$. By minimality of f_G we find $f_G \mid f$. The converse direction follows as $\mathbb{Q}(\zeta_{f_G}) \subset \mathbb{Q}(\zeta_f)$ if $f_G \mid f$.

So f_G is the greatest common divisor of all f, for which $L \subset \mathbb{Q}(\zeta_f)$. Equivalently f_G is the greatest common divisor of all $f \mid N$, for which $\ker((\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/f\mathbb{Z})^\times) \subset G$.

By class field theory, f_G is the finite part of the conductor of the abelian extension $L/\mathbb{Q}$.

The next lemma collects some classical facts on Gauss sums. We write $f_\chi = f_{\ker \chi}$ for a character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. We recall that $\boldsymbol{e}(\cdot)$ was defined in (1-3).

Lemma 3.1. *Let $N \in \mathbb{N}$ and say $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ is a character. For $k \in \mathbb{Z}$ we define $\tau = \sum_{\sigma \in (\mathbb{Z}/N\mathbb{Z})^\times} \chi(\sigma)\boldsymbol{e}(k\sigma/N)$, then the following hold true*:

(i) *If* $\gcd(k, N) = 1$ *then* $|\tau| \le f_\chi^{1/2}$.

(ii) *For unrestricted k we set $N' = N/\gcd(k, N)$. Then*

$$|\tau| \le \frac{\varphi(N)}{\varphi(N')} f_\chi^{1/2}.$$

Proof. If $k = 1$, part (i) follows directly from [Iwaniec and Kowalski 2004, Lemma 3.1, Section 3.4]. The more general case $\gcd(k, N) = 1$ follows as $\sum_{\sigma \in (\mathbb{Z}/N\mathbb{Z})^\times} \chi(\sigma)\boldsymbol{e}(k\sigma/N) = \sum_{\sigma \in (\mathbb{Z}/N\mathbb{Z})^\times} \chi(k'\sigma)\boldsymbol{e}(\sigma/N)$ where $kk' \equiv 1 \pmod N$ and since χ is completely multiplicative.

To prove (ii) set $N' = N/\gcd(k, N)$ and $k' = k/\gcd(k, N)$. Then τ is

$$\sum_{\sigma\in(\mathbb{Z}/N\mathbb{Z})^\times} \chi(\sigma)\boldsymbol{e}\left(\frac{k'}{N'}\sigma\right) = \sum_{\sigma'\in(\mathbb{Z}/N'\mathbb{Z})^\times} \Bigg(\sum_{\substack{\sigma\in(\mathbb{Z}/N\mathbb{Z})^\times \\ \sigma\equiv\sigma' \pmod{N'}}} \chi(\sigma)\Bigg)\boldsymbol{e}\left(\frac{k'}{N'}\sigma\right).$$

The inner sum on the right runs over a coset of the kernel of $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times$. Since χ is a character, the inner sum equals 0 if the said kernel does not lie in the kernel of χ. In this case, $\tau = 0$ and we are done.

Otherwise, $\ker((\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times) \subset \ker\chi$, and then $f_\chi \mid N'$. We find moreover that χ factors through a character $\chi' : (\mathbb{Z}/N'\mathbb{Z})^\times \to \mathbb{C}^\times$ and $f_{\chi'} \mid f_\chi$. As the kernel of $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times$ has order $\varphi(N)/\varphi(N')$ we have

$$\tau = \frac{\varphi(N)}{\varphi(N')} \sum_{\sigma'\in(\mathbb{Z}/N'\mathbb{Z})^\times} \chi'(\sigma')\boldsymbol{e}\left(\frac{k'}{N'}\sigma\right).$$

Part (ii) now follows from (i) since $\gcd(k', N') = 1$. □

Lemma 3.2. *Let $N \in \mathbb{N}$, let G be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$, and let $k \in \mathbb{Z}$. We define $N' = N/\gcd(k, N)$, then*

$$\frac{1}{\#G}\left|\sum_{\sigma\in G} \boldsymbol{e}(k\sigma/N)\right| \le \frac{[(\mathbb{Z}/N\mathbb{Z})^\times : G]}{\varphi(N')} f_G^{1/2}.$$

Proof. Let $\chi_1', \dots, \chi_m' : (\mathbb{Z}/N\mathbb{Z})^\times/G \to \mathbb{C}^\times$ be all characters and $m = [(\mathbb{Z}/N\mathbb{Z})^\times : G]$. Then $\sum_{i=1}^m \chi_i'(\sigma) = 0$ for all $\sigma \in (\mathbb{Z}/N\mathbb{Z})^\times/G$ expect for the neutral element, where this sum equals m. Write χ_i for $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N\mathbb{Z})^\times/G$ composed with χ_i'. Then $\sum_{i=1}^m \chi_i(\sigma) = 0$ if and only if $\sigma \in (\mathbb{Z}/N\mathbb{Z})^\times \backslash G$, otherwise this sum is m. Therefore,

$$\sum_{\sigma\in G} \boldsymbol{e}(k\sigma/N) = \frac{1}{m}\sum_{i=1}^m \sum_{\sigma\in(\mathbb{Z}/N\mathbb{Z})^\times} \chi_i(\sigma)\boldsymbol{e}(k\sigma/N) \tag{3-1}$$

and Lemma 3.1(ii) implies

$$\left|\sum_{\sigma\in(\mathbb{Z}/N\mathbb{Z})^\times} \chi_i(\sigma)\boldsymbol{e}(k\sigma/N)\right| \le \frac{\varphi(N)}{\varphi(N')} f_{\chi_i}^{1/2}.$$

Note that $G \subset \ker\chi_i$ because χ_i factors through $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N\mathbb{Z})^\times/G$. So $f_{\chi_i} \le f_G$, by the minimality of f_{χ_i}. The current lemma now follows from (3-1). □

Let $d, n \in \mathbb{N}$ and $x_1, \dots, x_n \in [0, 1)^d$. The *discrepancy* of $(x_1, \dots, x_n)$ is

$$\mathcal{D}(x_1, \dots, x_n) = \sup_B \left|\frac{\#\{i : x_i \in B\}}{n} - \mathrm{vol}(B)\right| \tag{3-2}$$

where B ranges over all products $\prod_{i=1}^d [\alpha_i, \beta_i)$ with $0 \le \alpha_i < \beta_i \le 1$. Note that the discrepancy lies in $[0, 1]$. In some references such as [Harman 1998], the discrepancy is not normalized by dividing by n and can be greater than 1.

In the next proposition we bound from above the discrepancy of the Galois orbit of a point of finite order in $\mathbb{G}_m^d$ using the Gauss sum estimates above. Below, $d_0(N)$ denotes the number of divisors of a natural number N.

Proposition 3.3. *Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have order N and let G be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ such that $\{\boldsymbol{\zeta}^\sigma : \sigma \in G\} = \{\boldsymbol{e}(x_i) : 1 \le i \le \#G\}$ with all x_i in $[0,1)^d$:*

(i) *We have*

$$\mathcal{D}(x_1, \dots, x_{\#G}) \ll_d [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{(\log 2\delta(\boldsymbol{\zeta}))^{d-1} \log\log 3\delta(\boldsymbol{\zeta})}{\delta(\boldsymbol{\zeta})^{1/2}}.$$

(ii) *If $d = 1$, then*

$$\mathcal{D}(x_1, \dots, x_{\#G}) \ll [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{\log(2N) d_0(N)}{\varphi(N)}.$$

Proof. We abbreviate $n = \#G$. We fix $a \in \mathbb{Z}^d$ with $\boldsymbol{\zeta} = \boldsymbol{e}(a/N)$. Then N and the entries of a are coprime. Let $H \ge 4$ be an integer. We use the Erdös–Turán–Koksma inequality [Harman 1998, Theorem 5.21], to bound the discrepancy $\mathcal{D} = \mathcal{D}(x_1, \dots, x_n)$ as follows

$$\mathcal{D} \ll_d \frac{1}{H} + \sum_{\substack{b \in \mathbb{Z}^d \setminus \{0\} \\ |b| \le H}} \frac{1}{r(b)} \left| \frac{1}{n} \sum_{\sigma \in G} \boldsymbol{e}\left(\frac{\langle a, b\rangle}{N} \sigma \right) \right| \tag{3-3}$$

here $r(b_1, \dots, b_d) = \max\{1, |b_1|\} \cdots \max\{1, |b_d|\}$.

By Lemma 3.2, the expression inside the modulus is at most $C/\varphi(N/\gcd(\langle a, b\rangle, N))$ with $C = [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2}$. We have $\varphi(M) \gg M/\log\log(3+M)$ for all integers $M \ge 1$ with an absolute and effective implicit constant, see for example [Rosser and Schoenfeld 1962, Theorem 15]. Therefore,

$$\mathcal{D} \ll_d \frac{1}{H} + C \sum_{\substack{b \in \mathbb{Z}^d \setminus \{0\} \\ |b| \le H}} \frac{1}{r(b)} \frac{\gcd(\langle a, b\rangle, N)}{N} \log\log(3 + N/\gcd(\langle a, b\rangle, N)).$$

If $b \in \mathbb{Z}^d \setminus \{0\}$, then

$$\left\langle a, \frac{N}{\gcd(\langle a, b\rangle, N)} b \right\rangle = N \frac{\langle a, b\rangle}{\gcd(\langle a, b\rangle, N)} \in N\mathbb{Z}$$

which implies $\boldsymbol{\zeta}^{bN/\gcd(\langle a,b\rangle,N)} = 1$. So $N/\gcd(\langle a, b\rangle, N) \ge \delta/|b| > 0$ where $\delta = \delta(\boldsymbol{\zeta})$. As $t \mapsto (\log\log(3+t))/t$ is decreasing on $t > 0$ we find

$$\mathcal{D} \ll_d \frac{1}{H} + C \frac{1}{\delta} \sum_{\substack{b \in \mathbb{Z}^d \setminus \{0\} \\ |b| \le H}} \frac{|b|}{r(b)} \log\log(3 + \delta).$$

The sum of $|b|/r(b)$ over all $b \in \mathbb{Z}^d$ with $1 \le |b| \le H$ is $\ll_d H(\log H)^{d-1}$, so we find

$$\mathcal{D} \ll_d \frac{1}{H} + C \frac{\log\log(3\delta)}{\delta} H(\log H)^{d-1}.$$

Part (i) follows by fixing H to be the least integer with $H \ge \delta^{1/2}$ and $H \ge 4$.

In part (ii) we have $d = 1$. We may assume $N \geq 4$ as the discrepancy is at most 1. Here a is coprime to N and so $\gcd(ab, N) = \gcd(b, N)$. In (3-3) we take $H = N$ and use again Lemma 3.2 with C as before to find

$$\mathcal{D} \ll \frac{1}{N} + \sum_{b=1}^{N} \frac{C}{b\varphi(N/\gcd(b,N))} \leq \frac{1}{N} + \sum_{g|N} \frac{C}{g\varphi(N/g)} \sum_{e=1}^{N/g} \frac{1}{e}.$$

In the sum over g we have $g\varphi(N/g) \geq \varphi(N)$ and the harmonic sum is $\ll \log N$. So $\mathcal{D} \ll 1/N + C(\log N)d_0(N)/\varphi(N)$, which implies (ii). □

A variant of the case $d = 1$ already appears in [Baker et al. 2008, Lemma 1.3] which is attributed to Pomerance.

The discrepancy bound in (i) depends on $\delta(\boldsymbol{\zeta})$. But $\delta(\boldsymbol{\zeta})$ is always bounded above by N. So estimates that decay in N are stronger than estimates that decay in $\delta(\boldsymbol{\zeta})$. However, there can be no upper bound for the discrepancy in terms of the order N.

Let us assume for the moment $d = 1$. Then we have $\delta(\boldsymbol{\zeta}) = N$. If $[(\mathbb{Z}/N\mathbb{Z})^\times : G]$ and f_G are fixed, the decay of the discrepancy is $1/N$ up to terms of subpolynomial growth by part (ii) of the preceding proposition and standard estimates for Euler's function φ.

The *total variation* on $[a, b]$ of a real valued function F whose domain contains the interval $[a, b]$ with $a \leq b$ is

$$\mathrm{Var}_a^b(F) = \sup_{a \leq x_0 \leq \cdots \leq x_m \leq b} \sum_{i=1}^{m} |F(x_i) - F(x_{i-1})|.$$

For $a = 0$ and $b = 1$ we abbreviate $\mathrm{Var}(F) = \mathrm{Var}_a^b(F)$.

The next lemma requires Koksma's inequality.

Lemma 3.4. *Let $F : [0, 1] \to \mathbb{R}$ be a function with $\mathrm{Var}(F) < \infty$. If $N \geq 1$ is an integer and G is a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ such that $\{\zeta^\sigma : \sigma \in G\} = \{\boldsymbol{e}(x_i) : 1 \leq i \leq \#G\}$ with all x_i in $[0, 1)$, then*

$$\left| \frac{1}{\#G} \sum_{i=1}^{\#G} F(x_i) - \int_0^1 F(x)\,dx \right| \ll [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{\log(2N)d_0(N)}{\varphi(N)} \mathrm{Var}(F).$$

Proof. The claim follows from Theorems 1.3 and 5.1 in Chapter 2 of [Kuipers and Niederreiter 1974] together with Proposition 3.3(ii). □

3A. *A univariate average.*

Lemma 3.5. *Let $\alpha \in \mathbb{C}$ and $r > 0$. For $x \in [0, 1]$ we define*

$$F_{\alpha,r}(x) = \log \max(r, |\boldsymbol{e}(x) - \alpha|).$$

Then $F_{\alpha,r} : [0, 1] \to \mathbb{R}$ satisfies $\mathrm{Var}(F_{\alpha,r}) \leq 3 \log(1 + 2/r)$.

Proof. We abbreviate $F = F_{\alpha,r}$. By elementary geometry we can find $m \leq 3$ and $0 = x_0 < x_1 < \cdots < x_m = 1$ such that F is monotone on all $[x_{i-1}, x_i]$. Then $\mathrm{Var}_{x_{i-1}}^{x_i}(F) = |F(x_i) - F(x_{i-1})|$ and

$\mathrm{Var}(F) = \sum_{i=1}^{m} \mathrm{Var}_{x_{i-1}}^{x_i}(F)$. We have $\log\max\{r, |\alpha|-1\} \le F(x) \le \log\max\{r, |\alpha|+1\}$ for all $x \in [0,1]$. Hence $\mathrm{Var}_{x_{i-1}}^{x_i}(F) \le \log(\max\{r, |\alpha|+1\}/\max\{r, |\alpha|-1\})$ which we see is at most $\log(1+2/r)$ by considering the cases $|\alpha| \ge 1+r$ and $|\alpha| < 1+r$. Thus $\mathrm{Var}(F) \le 3\log(1+2/r)$. □

The value r serves as a truncation parameter. We now apply Koksma's inequality to $F_{\alpha,r}$.

Lemma 3.6. *Let $\zeta \in \mu_\infty$ have order N and let G be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$. Let $\alpha \in \mathbb{C}$ and $r \in (0,1]$, then*

$$\frac{1}{\#G} \sum_{\substack{\sigma \in G \\ |\zeta^\sigma - \alpha| > r}} \log|\zeta^\sigma - \alpha| = \log^+|\alpha| + O\left(\left([(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{\log(2N) d_0(N)}{\varphi(N)} + r\right)\left|\log\frac{r}{2}\right|\right). \tag{3-4}$$

Proof. We let I denote the left-hand side of (3-4). Then $I = I_1 + I_2$ with

$$I_1 = \frac{1}{\#G}\sum_{i=1}^{\#G} F_{\alpha,r}(x_i) \quad \text{and} \quad I_2 = \frac{1}{\#G} \sum_{\substack{i \\ |\boldsymbol{e}(x_i)-\alpha| \le r}} -\log r$$

with the $x_i \in [0,1)$ as in Lemma 3.4 and $F_{\alpha,r}$ as in Lemma 3.5. The integrals below are understood to be over subsets of $[0,1]$. Applying Lemmas 3.4 and 3.5 to $F_{\alpha,r}$ yields

$$I_1 = \int_0^1 F_{\alpha,r}(x)\,dx + O\left([(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{\log(2N) d_0(N)}{\varphi(N)}\left|\log\frac{r}{2}\right|\right). \tag{3-5}$$

The set of $x \in [0,1]$ with $|\boldsymbol{e}(x) - \alpha| \le r$ is of the form $\varnothing$, $[a,b]$, or $[0,a]\cup[b,1]$. So its characteristic function has total variation at most 2. Lemma 3.4 applied to this characteristic function yields

$$I_2 = -\int_{|\boldsymbol{e}(x)-\alpha|\le r} \log r\,dx + O\left([(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \frac{\log(2N) d_0(N)}{\varphi(N)}\right). \tag{3-6}$$

The sum of the integrals in (3-5) and (3-6) equals

$$\int_0^1 \log|\boldsymbol{e}(x) - \alpha|\,dx - \int_{|\boldsymbol{e}(x)-\alpha|\le r} \log|\boldsymbol{e}(x)-\alpha|\,dx.$$

Jensen's formula [Bombieri and Gubler 2006, Proposition 1.6.5] implies that the first integral equals $\log^+|\alpha|$. To complete the proof it suffices to show that the second integral is $O(r|\log r/2|)$.

The integral is nonpositive as $r \le 1$ and we may assume that it is nonzero. First assume, $|\alpha| \le \frac{1}{2}$. In this case $|\boldsymbol{e}(x) - \alpha| \ge \frac{1}{2}$ and the integral is $O(r)$. Second, say $|\alpha| > \frac{1}{2}$. [Rahman and Schmeisser 2002, Lemma 11.6.1] implies $|\boldsymbol{e}(x) - \alpha| \ge |\alpha|^{1/2}|\boldsymbol{e}(x) - \boldsymbol{e}(y)| \ge 2^{-1/2}|\boldsymbol{e}(x-y) - 1|$ where $\alpha = |\alpha|\boldsymbol{e}(y)$ and $|x-y| \le \frac{1}{2}$. There is an absolute and effectively computable constant $C > 0$ with $|\boldsymbol{e}(x-y) - 1| \ge C|x-y|$ and thus $|\boldsymbol{e}(x) - \alpha| \ge 2^{-1/2}C|x-y|$. In the integral we have $r \ge |\boldsymbol{e}(x) - \alpha|$ and so the desired bound follows from elementary analysis. □

3B. ***A Galois conjugate near*** **1.** We will also need an estimate on the minimal distance of a Galois conjugate of a torsion point to the unit element.

Lemma 3.7. *Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have order N and let G be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$. There exist $\sigma \in G$ and $a \in \mathbb{Z}^d$ with $\boldsymbol{\zeta} = \boldsymbol{e}(a\sigma/N)$, $|a| < N$, and*

$$\frac{|a|}{N} \ll_d \frac{[(\mathbb{Z}/N\mathbb{Z})^\times : G]^{1/d} f_G^{1/(2d)}}{\delta(\boldsymbol{\zeta})^{1/(3d)}}. \tag{3-7}$$

Proof. Let $\boldsymbol{\zeta} = \boldsymbol{e}(b/N)$ with $b \in \mathbb{Z}^d$, the entries of b and N have no common prime divisor. Suppose $x_1, \dots, x_n$ are as in Proposition 3.3 coming from the $\boldsymbol{\zeta}^\sigma$ as σ ranges over G where $n = \#G$. There exists $c(d) > 0$ depending only on d with $\mathcal{D}(x_1, \dots, x_n) \le c(d)[(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \delta(\boldsymbol{\zeta})^{-1/3}$; note that $\delta(\boldsymbol{\zeta}) = N$ if $d = 1$. We set $\kappa = 2c(d)^{1/d}[(\mathbb{Z}/N\mathbb{Z})^\times : G]^{1/d} f_G^{1/(2d)} \delta(\boldsymbol{\zeta})^{-1/(3d)}$. If $\kappa \ge 1$ we take σ the identity and fix $a \in \mathbb{Z}^d$ with $|a| < N$ and $\boldsymbol{\zeta} = \boldsymbol{e}(a/N)$. Otherwise, by the definition of the discrepancy the hypercube $[0, \kappa)^d$ contains some $x_i = a/N$. Hence a satisfies, $|a| < N$, (3-7), and $\boldsymbol{e}(a/N) = \boldsymbol{\zeta}^{\sigma^{-1}}$ for some $\sigma \in G$. □

4. Theorem of Mahler and Mignotte

In this section, we establish the separation of pairs of roots of an integer polynomial. Theorem 4.1 below was shown by Mahler [1964] for the case $k = 1$ of a single pair of roots. Mignotte [1995] generalized Mahler's inequality to products over several disjoint pairs of roots (see his Theorem 1). We reproduce here a lightened version of Mignotte's theorem that is suitable for our needs. The proof is an adaptation of Mahler's original argument about a single pair, guided by the principle that Liouville's Inequality bounds an algebraic number at an arbitrary set of places in terms of the height. Let us also mention Güting's proof [1961] of a less precise earlier result involving the length of a polynomial instead of the Mahler measure.

Let $Q \in \mathbb{C}[X]$ be a nonzero univariate polynomial. By Jensen's formula its Mahler measure equals

$$m(Q) = \log|a_0| + \sum_{i=1}^{D} \log^+|z_i| \tag{4-1}$$

if $Q = a_0(X - z_1)\cdots(X - z_D)$ and where the z_i are complex. If Q is nonconstant, we let $\operatorname{disc}(Q)$ denote its discriminant as a degree $\deg Q$ polynomial.

Theorem 4.1. *Let $Q \in \mathbb{C}[X] \setminus \mathbb{C}$ be of degree D and with no repeated roots. If $z_1, \dots, z_k, z_1', \dots, z_k'$ are pairwise distinct complex roots of Q, then*

$$\sum_{j=1}^{k} -\log|z_j - z_j'| \le \tfrac{1}{2}(D + 2k)\log D - \tfrac{k}{2}\log 3 + (D-1)m(Q) - \tfrac{1}{2}\log|\operatorname{disc}(Q)| \tag{4-2}$$

with strict inequality for $k \ge 1$.

Proof. We modify Mahler's and Mignotte's argument as follows.

Both sides of (4-2) are invariant under multiplication Q by a nonzero scalar. So we may assume that Q is monic. After possibly swapping z_j with z'_j we may assume $|z_j| \geq |z'_j|$ for all j.

We augment $z_1, \dots, z_k$ to all complex roots $z_1, \dots, z_D$ of Q. Then we consider the Vandermonde determinant

$$V = \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_D \\ \vdots & \vdots & & \vdots \\ z_1^{D-1} & z_2^{D-1} & \cdots & z_D^{D-1} \end{pmatrix},$$

which is nonzero as $z_1, \dots, z_D$ are pairwise distinct. For $j \in \{1, \dots, k\}$, let $i_j > k$ be the index with $z'_j = z_{i_j}$. For these j, we subtract the i_j-th column from the j-th column and factoring each difference $z_j - z_{i_j}$ out of the determinant with the identities $z_j^m - z_{i_j}^m = (z_j - z_{i_j})(z_j^{m-1} + z_j^{m-2} z_{i_j} + \cdots + z_{i_j}^{m-1})$, $1 \leq m \leq D-1$. We obtain an expression

$$V = W \prod_{j=1}^{k} (z_j - z_{i_j}) = W \prod_{j=1}^{k} (z_j - z'_j), \tag{4-3}$$

where $W \neq 0$ is the determinant of the matrix having

$$\begin{pmatrix} 0 \\ 1 \\ z_j + z'_j \\ \vdots \\ z_j^{D-2} + z_j^{D-3} z'_j + \cdots + {z'_j}^{D-2} \end{pmatrix}$$

for its j-th column, $j \in \{1, \dots, k\}$, and the same entries as in the Vandermonde matrix in the remaining columns. By Hadamard's inequality, $|W|$ is bounded from above by the product of the Hermitian norms of all these columns. The j-th column, for some $j \in \{1, \dots, k\}$, has Hermitian norm

$$\sqrt{\sum_{m=0}^{D-2} |z_j^m + z_j^{m-1} z'_j + \cdots + {z'_j}^m|^2} \leq \sqrt{\sum_{m=0}^{D-2} (m+1)^2} \max\{1, |z_j|, |z'_j|\}^{D-2} < \sqrt{D^3/3} \cdot \max\{1, |z_j|\}^{D-1}$$

where we used $|z'_j| \leq |z_j|$. The Hermitian norm of the j-th column with $j \in \{k+1, \dots, D\}$ is at most $\sqrt{D} \max\{1, |z_j|\}^{D-1}$.

Applying Hadamard's inequality, using these two bounds, and taking the logarithm yields

$$\begin{aligned} \log|W| &\leq \frac{k}{2} \log\left(\frac{D^3}{3}\right) + \frac{D-k}{2} \log D + (D-1) \sum_{j=1}^{D} \log^+ |z_j| \\ &= \frac{D+2k}{2} \log D - \frac{k}{2} \log 3 + (D-1) m(Q) \end{aligned}$$

as Q is monic; the inequality is strict for $k \geq 1$. If $k = 0$ no column operations are necessary and $V = W$.

The monic polynomial Q has discriminant $\operatorname{disc}(Q) = V^2$. Consequently $|V| = |\operatorname{disc}(Q)|^{1/2}$, and in view of (4-3) we have

$$\sum_{j=1}^{k} -\log|z_j - z_j'| = \log|W| - \log|V| \le \tfrac{1}{2}(D+2k)\log D - \tfrac{k}{2}\log 3 + (D-1)m(Q) - \tfrac{1}{2}\log|\operatorname{disc}(Q)|$$

with a strict inequality for $k \ge 1$. This concludes the proof. □

While Theorem 4.1 suffices for our needs here, we remark that it is possible to relax the hypothesis to having $z_1, \dots, z_k$ pairwise distinct and $\{z_1, \dots, z_k\} \cap \{z_1', \dots, z_k'\} = \varnothing$, at the cost of a slightly worse upper bound (4-2).

The following corollary holds for integral polynomials that are not necessarily squarefree.

Corollary 4.2. *Let $Q \in \mathbb{Z}[X] \setminus \mathbb{Z}$ be of degree D. If $z_1, \dots, z_k, z_1', \dots, z_k'$ are pairwise distinct complex roots of Q, then*

$$\sum_{j=1}^{k} -\log|z_j - z_j'| \le \frac{D+2k}{2}\log D - \frac{k}{2}\log 3 + (D-1)m(Q) \tag{4-4}$$

with strict inequality for $k \ge 1$.

Proof. We may assume $k \ge 1$. We begin by splitting off the squarefree part of Q. More precisely, we factor $Q = \widetilde{Q}R$ where $\widetilde{Q}, R \in \mathbb{Z}[X]$ and $\widetilde{Q}$ is squarefree and vanishes at all complex roots of Q. The discriminant $\operatorname{disc}(\widetilde{Q})$ is a nonzero integer, and so $|\operatorname{disc}(\widetilde{Q})| \ge 1$. Moreover, $m(\widetilde{Q}) \ge 0$. Theorem 4.1 applied to $\widetilde{Q}$ and $1 \le \deg \widetilde{Q} \le D$ implies that the sum on the left of (4-4) is at most $\frac{1}{2}(D+2k)\log D - \frac{k}{2}\log 3 + (D-1)m(\widetilde{Q})$. The corollary follows from $m(\widetilde{Q}) = m(Q) - m(R) \le m(Q)$. □

4A. ***A repulsion property of the unit circle.*** A key point in [Habegger 2018] is that while Mahler's theorem does not give a strong enough bound for the distance of a complex root of $Q \in \mathbb{Z}[X] \setminus \{0\}$ to an N-th root of unity (the product $(X^N - 1)Q(X)$ has an exceedingly large degree), it can be used to bound the distance from the unit circle to the locus of roots of P lying off the unit circle. With Corollary 4.2, this repulsion property of the unit circle can be strengthened as follows.

Lemma 4.3. *Let $Q \in \mathbb{Z}[X] \setminus \mathbb{Z}$ and $Q = a_0(X - z_1)\cdots(X - z_D)$ where $z_1, \dots, z_D \in \mathbb{C}$. Then*

$$\sum_{\substack{j=1 \\ |z_j| \ne 1}}^{D} \log^+ \frac{1}{\big||z_j| - 1\big|} \le D\log\left(\frac{3+\sqrt{5}}{2}\right) + 2D\log(2D) + 4Dm(Q) \le 4D(\log(2D) + m(Q)). \tag{4-5}$$

Before we come to the proof let us remark that $||z| - 1|$ is the distance $\operatorname{dist}(z, S^1)$ of $z \in \mathbb{C}$ to the unit circle S^1. Thus inequality (4-5) can be restated as providing

$$\frac{1}{D}\sum_{\substack{j=1 \\ |z_j| \ne 1}}^{D} \log^+ \frac{1}{\operatorname{dist}(z_j, S^1)} \le \log\left(\frac{3+\sqrt{5}}{2}\right) + 2\log(2D) + 4m(Q).$$

Our result suggests that the unit circle repels roots of Q that lie off the unit circle. Related estimates are implicit in work of Dubickas [1997], see his Theorem 2.

Proof. The second bound in (4-5) is elementary, so it suffices to prove the first one.

Say $Q = a_0 Q_1 \cdots Q_n$ where each $Q_i \in \mathbb{Z}[X]$ is irreducible of positive degree with $a_0 \in \mathbb{Z}$. Observe that $0 \le m(Q_i) \le m(Q)$ and $\sum_i \deg Q_i = \deg Q$. So it suffices to prove (4-5) for Q irreducible in $\mathbb{Z}[X]$. We may also assume $Q(0) \neq 0$.

We will apply Corollary 4.2 to the polynomial $\widetilde{Q} \in \mathbb{Z}[X]$ constructed from Q in the following manner. If $Q(1/X)X^D \neq \pm Q$ we take $\widetilde{Q} = Q(X)Q(1/X)X^D$ and $\widetilde{Q} = Q$ otherwise. So $\widetilde{D} = \deg \widetilde{Q} = \delta D$ and $m(\widetilde{Q}) = \delta m(Q)$ with $\delta = 2$ in the first case and $\delta = 1$ in the second case. Indeed, a polynomial and its reciprocal have the same Mahler measure. For any root z of $\widetilde{Q}$ we also have $\widetilde{Q}(1/\bar{z}) = 0$.

The following basic observation for a complex number z will prove useful. We have $|z - 1/\bar{z}| \le 1$ if and only if $\phi^{-1} \le |z| \le \phi$ with $\phi = (1+\sqrt{5})/2$ the golden ratio.

Let $w_1, \dots, w_k$ be the roots of $\widetilde{Q}$ without repetition such that $\phi^{-1} \le |w_j| < 1$. Then $w_j' = 1/\overline{w}_j$ is a root of $\widetilde{Q}$ for each $j \in \{1, \dots, k\}$ with $|w_j'| > 1$. Corollary 4.2 yields

$$\sum_{j=1}^{k} \log^+ \frac{1}{|w_j - 1/\overline{w}_j|} \le \delta D \log(\delta D) + \delta^2 D m(Q) \tag{4-6}$$

because $k \le \widetilde{D}/2 = \delta D/2$ and $m(Q) \ge 0$.

Suppose z_j is a root of Q with $|z_j| \neq 1$ and $\phi^{-1} \le |z_j| \le \phi$. Then $z_j \in \{w_l, 1/\overline{w}_l\}$ for some unique l. The mapping $j \mapsto l$ is at worst 2-to-1 and injective if $\delta = 2$ as Q is irreducible.[1] This leads to the factor $2/\delta$ in

$$\sum_{\substack{|z_j|\neq 1 \\ 1/\phi \le |z_j| \le \phi}} \log^+ \frac{1}{|z_j - 1/\bar{z}_j|} \le \frac{2}{\delta} \sum_{l=1}^{k} \log^+ \frac{1}{|w_l - 1/\overline{w}_l|} \tag{4-7}$$

For a complex number z with $|z| \ge \phi^{-1}$ we have $|z - 1/\bar{z}| = \frac{|z|+1}{|z|}\big||z| - 1\big| \le (1+\phi)\big||z| - 1\big|$. This allows us to get

$$\sum_{\substack{|z_j|\neq 1 \\ 1/\phi \le |z_j| \le \phi}} \log^+ \frac{1}{\big||z_j| - 1\big|} \le s \log(1+\phi) + \sum_{\substack{|z_j|\neq 1 \\ 1/\phi \le |z_j| \le \phi}} \log^+ \frac{1}{|z_j - 1/\bar{z}_j|}$$

where s is the number of terms in the first sum. There are $D - s$ remaining roots of Q and if $|z_j| < \phi^{-1}$ or $|z_j| > \phi$ we get $\log^+ 1/\big||z| - 1\big| \le \log(1+\phi)$. Together with (4-6) and (4-7) we find

$$\sum_{|z_j|\neq 1} \log^+ \frac{1}{\big||z_j| - 1\big|} \le D \log(1+\phi) + 2D \log(\delta D) + 2\delta D m(Q).$$

We have established (4-5) for Q as $\delta \le 2$. □

Next we generalize our bound to a polynomial with coefficients in a number field. Recall that $h(Q)$ is the absolute logarithmic projective Weil height of a nonzero polynomial Q with algebraic coefficients.

[1] Indeed, if $z_j, z_k \in \{w_l, 1/\overline{w}_l\}$ with $z_j \neq z_k$, then $z_j = 1/\bar{z}_k$. So $\widetilde{Q} = Q$ and hence $\delta = 1$ in this case.

Corollary 4.4. *Let $F \subset \mathbb{C}$ be a number field and let $Q \in F[X] \setminus F$ and $Q = a_0(X - z_1)\cdots(X - z_D)$ where $z_1, \dots, z_D \in \mathbb{C}$. Then*

$$\sum_{\substack{j=1 \\ |z_j|\neq 1}}^{D} \log^+ \frac{1}{\big||z_j| - 1\big|} \le 10[F : \mathbb{Q}]^2 D(\log(2D) + h(Q)).$$

Proof. Let $\widetilde{Q}$ be the product of the $\mathbb{Q}$-Galois conjugates of Q. Then $\widetilde{Q}$ has rational coefficients and degree $\widetilde{D} \le D[F : \mathbb{Q}]$. Let $\lambda \in \mathbb{N}$ such that $\lambda\widetilde{Q}$ is integral with content 1. For the projective height we find $h(\widetilde{Q}) = \log|\lambda\widetilde{Q}|$. Together with [Bombieri and Gubler 2006, Lemma 1.6.7] we get $m(\lambda\widetilde{Q}) \le \frac{1}{2}\log(1+\widetilde{D}) + h(\widetilde{Q})$. As all $\mathbb{Q}$-Galois conjugates of Q have the same projective height we use elementary estimates at local places, see [Bombieri and Gubler 2006, Remark 1.6.14], to find

$$h(\widetilde{Q}) \le [F : \mathbb{Q}]\log(1 + D) + [F : \mathbb{Q}]h(Q).$$

By Lemma 4.3 applied to $\lambda\widetilde{Q}$, the sum $\sum_{j=1:|z_j|\neq 1}^{D} \log^+ 1/\big||z_j| - 1\big|$ is at most

$$4\widetilde{D}\big(\log(2\widetilde{D}) + \tfrac{1}{2}\log(1 + \widetilde{D}) + [F : \mathbb{Q}]\log(1 + D) + [F : \mathbb{Q}]h(Q)\big).$$

We use $1 + \widetilde{D} \le 2\widetilde{D} \le 2D[F : \mathbb{Q}] \le (2D)^{[F:\mathbb{Q}]}$ to complete the proof. □

4B. ***Averages over roots of unity.*** In this subsection we apply the repulsion property of the unit circle, Corollary 4.4, to estimate the norm of cyclotomic integers of the form $Q(\zeta)$, where ζ is a varying root of unity and Q is a moderately controlled univariate polynomial with algebraic coefficients and without zeros in $S^1 \setminus \mu_\infty$. This gives a fairly uniform solution of the one dimensional essentially atoral case and forms the basis for the higher dimensional case to be taken up in the next sections.

Proposition 4.5. *Let $F \subset \mathbb{C}$ be a number field and let $Q \in F[X]\setminus\{0\}$ be of degree at most $D \ge 1$ with no roots in $S^1\setminus\mu_\infty$. Let $\zeta \in \mu_\infty$ be of order N and G a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ such that $Q(\zeta^\sigma) \neq 0$ for all $\sigma \in G$. Then*

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|Q(\zeta^\sigma)| = m(Q) + O\left([F : \mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} D(\log(2D) + h(Q))\frac{(\log 2N)^3 d_0(N)}{N}\right). \quad (4\text{-}8)$$

Proof. We may assume that Q is nonconstant and $D = \deg Q$. Let $Q = a_0(X - z_1)\cdots(X - z_D)$. The idea is that each given root z_j may get within distance of $\le 1/N^2$ to at most a single conjugate of ζ.

We call z_j exceptional if $|\zeta^{\sigma_j} - z_j| \le 1/N^2$ for some $\sigma_j \in G$. As $|\xi - \xi'| \ge 4/N$ for distinct roots of unity ξ, ξ' of order N we see that σ_j is uniquely determined by z_j. Note that $\zeta^{\sigma_j} \neq z_j$ because $Q(\zeta^{\sigma_j}) \neq 0 = Q(z_j)$.

We apply Lemma 3.6 with $\alpha = z_j$ and $r = 1/N^2$. Thus

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|\zeta^\sigma - z_j| = \log^+|z_j| + \frac{1}{\#G}\log|\zeta^{\sigma_j} - z_j| + O\left([(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2}\frac{(\log 2N)^2 d_0(N)}{\varphi(N)} + \frac{\log 2N}{N^2}\right), \tag{4-9}$$

if z_j is exceptional, otherwise the same bound without the term $(\#G)^{-1}\log|\zeta^{\sigma_j} - z_j|$ holds true. As $1/N^2 \le 1/\varphi(N)$ we merge $(\log 2N)/N^2$ into the first term of the error term. Summing (4-9) over all $j \in \{1, \dots, D\}$ and adding $\log|a_0|$ gives

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|Q(\zeta^\sigma)| = m(Q) - \frac{1}{\#G}\sum_{j=1}^{D}{}' \log\frac{1}{|\zeta^{\sigma_j} - z_j|} + O\left([(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} D\frac{(\log 2N)^2 d_0(N)}{\varphi(N)}\right), \tag{4-10}$$

the dash signifies that we only sum over those j for which z_j is exceptional.

To bound the dashed sum we require Corollary 4.4. If z_j is exceptional, then $|\zeta^{\sigma_j} - z_j| \le 1$. Therefore, the dashed sum is nonnegative.

First, we consider the subsum over all exceptional $z_j \notin \mu_\infty$. Then $|z_j| \ne 1$ and $|\zeta^{\sigma_j} - z_j| \ge ||z_j| - 1|$ by the reverse triangle inequality. By Corollary 4.4 we find

$$0 \le \sum_{\substack{j=1\\ z_j\notin\mu_\infty}}^{D}{}' \log\frac{1}{|\zeta^{\sigma_j} - z_j|} \le \sum_{\substack{j=1\\ z_j\notin\mu_\infty}}^{D}{}' \log^+\frac{1}{||z_j| - 1|} = O([F:\mathbb{Q}]^2 D(\log(2D) + h(Q))). \tag{4-11}$$

Second, we consider the subsum over all exception $z_j \in \mu_\infty$, which is harmless. Recall that $\zeta^{\sigma_j} \ne z_j$. Since the order of z_j is $\ll [\mathbb{Q}(z_j):\mathbb{Q}]^2 \le (D[F:\mathbb{Q}])^2$ and the order of ζ^{σ_j} is N we find $|\zeta^{\sigma_j} - z_j| \gg N^{-1}(D[F:\mathbb{Q}])^{-2}$. On the other hand, $|\zeta^{\sigma_j} - z_j| \le N^{-2}$ and hence $N \ll (D[F:\mathbb{Q}])^2$. We obtain the crude estimate $|\zeta^{\sigma_j} - z_j| \gg (D[F:\mathbb{Q}])^{-4} \gg (2D)^{-4[F:\mathbb{Q}]}$ and finally bound the at most D terms below separately to get

$$0 \le \sum_{\substack{j=1\\ z_j\in\mu_\infty}}^{D}{}' \log\frac{1}{|\zeta^{\sigma_j} - z_j|} = O([F:\mathbb{Q}]D\log(2D)). \tag{4-12}$$

We divide the sum of (4-11) and (4-12) by $\#G$ to find

$$0 \le \frac{1}{\#G}\sum_{j=1}^{D}{}' \log\frac{1}{|\zeta^{\sigma_j} - z_j|} = O\left([F:\mathbb{Q}]^2 D(\log(2D) + h(Q))\frac{[(\mathbb{Z}/N\mathbb{Z})^\times : G]}{\varphi(N)}\right).$$

The proposition follows from (4-10) and $\varphi(N) \gg N/\log\log(3N)$, a consequence of [Rosser and Schoenfeld 1962, Theorem 15]. □

Proposition 4.5 and ultimately Theorem 4.1 may be viewed as our input from transcendence theory. If this or a comparable bound held without the restrictive condition that Q has no roots in $S^1 \setminus \mu_\infty$ then it could be used to attack Conjecture 1.3. We were unable to prove or disprove that a suitable version of Proposition 4.5 extends to general polynomials. Progress on Conjecture 1.5 could indicate a path towards this goal.

5. Geometry of numbers

Let $d \geq 1$ and suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has order N. It would be useful if $\boldsymbol{\zeta}$ had a Galois conjugate close to the unit element 1. If the distance were at most a small power of N^{-1}, this conjugate could be used to help reduce the multivariate Theorem 1.1 to the univariate Proposition 4.5, see [Habegger 2018].

Unfortunately, such a conjugate need not exist. Take for example $\boldsymbol{\zeta} = \boldsymbol{e}(1/p, 1/p^n)$ where p is a prime and $n \in \mathbb{N}$, here $N = p^n$. Any conjugate of $\boldsymbol{\zeta}$ has distance $\gg 1/p$ to 1 regardless of the value of n. The problem is that $\boldsymbol{\zeta}$ is up to a point of order p contained in the algebraic subgroup $\{1\} \times \mathbb{G}_m$.

We overcome this difficult by constructing a factorization $\boldsymbol{\zeta} = \boldsymbol{\eta}\boldsymbol{\xi}$ into torsion points $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ that satisfy the following properties for prescribed $\epsilon > 0$. First, the order of $\boldsymbol{\eta}$ is small relative to N, more precisely it is $O_{d,\epsilon}(N^\epsilon)$. Second, some Galois conjugate of $\boldsymbol{\xi}$ is at distance at most $O_{d,\epsilon}(N^{-\kappa(\epsilon)})$ to 1. Here $\kappa(\epsilon)$ is expected to be small for small ϵ. But we will see that $\kappa(\epsilon)/\epsilon$ is large. This is of central importance for our application.

We use the geometry of numbers to construct this factorization. An important tool is the slope of a lattice.

A lattice Λ in $\mathbb{R}^d$ is a finitely generated and discrete subgroup of $\mathbb{R}^d$. The rank of Λ is denoted by $\operatorname{rk}(\Lambda)$ and its determinant by $\det(\Lambda)$. We consider the set

$$A = \{(r, \log\det(\Omega)) : r \in \mathbb{Z} \text{ and } \Omega \text{ is a subgroup of } \Lambda \text{ with } \operatorname{rk}(\Omega) = r\}$$

and use the convention $\det(\{0\}) = 1$. In contrast to the convention in Arakelov theory, we have no sign in front of $\log\det(\Lambda)$. Observe that the second coordinate is bounded from below on A. Stuhler [1976, Proposition 1] proved that for each $j \in \{0, \dots, \operatorname{rk}(\Lambda)\}$ there exists a sublattice $\Lambda_j \subset \Lambda$ of rank j, possibly nonunique, with $\log\det(\Lambda_j)$ minimal among all sublattices of rank j. The lower boundary of the convex hull of A is the graph of a piecewise linear, continuous, convex function $P : [0, \operatorname{rk}(\Lambda)] \to \mathbb{R}$. As $\Lambda_0 = \{0\}$ and $\Lambda_{\operatorname{rk}(\Lambda)} = \Lambda$ we find $P(0) = 0$ and $P(\operatorname{rk}(\Lambda)) = \log\det(\Lambda)$.

For each $j \in \{1, \dots, \operatorname{rk}(\Lambda)\}$, the slope of P on $[j-1, j]$ is

$$\mu_j(\Lambda) = P(j) - P(j-1).$$

By convexity we have

$$\mu_1(\Lambda) \leq \mu_2(\Lambda) \leq \cdots \leq \mu_{\operatorname{rk}(\Lambda)}(\Lambda).$$

Moreover, $\mu_1(\Lambda) + \cdots + \mu_j(\Lambda) = P(j) - P(0) = P(j)$ for all j as $P(0) = \log\det(\Lambda_0) = 0$.

Assume $\Lambda \neq \{0\}$ and let $\nu \in \left(0, \frac{1}{2}\right]$ be a parameter. Suppose that

$$\mu_j(\Lambda) < \nu^{\mathrm{rk}(\Lambda)-j+1} \log \det(\Lambda)$$

for all $j \in \{1, \dots, \mathrm{rk}(\Lambda)\}$. Taking the sum yields

$$\log \det(\Lambda) < (\nu + \nu^2 + \cdots + \nu^{\mathrm{rk}(\Lambda)}) \log \det(\Lambda).$$

As $\nu \in \left(0, \frac{1}{2}\right]$ we must have $\det(\Lambda) < 1$.

Let us now assume $\det(\Lambda) \geq 1$, then there exists a unique $j_0 \in \{0, \dots, \mathrm{rk}(\Lambda) - 1\}$ such that

$$\mu_k(\Lambda) < \nu^{\mathrm{rk}(\Lambda)-k+1} \log \det(\Lambda) \quad \text{for all } 1 \leq k \leq j_0 \text{ and} \quad \mu_{j_0+1}(\Lambda) \geq \nu^{\mathrm{rk}(\Lambda)-j_0} \log \det(\Lambda). \tag{5-1}$$

We write $\Lambda(\nu)$ for the rank j_0 lattice Λ_{j_0}, indicating its dependency on ν. It satisfies $\mathrm{rk}(\Lambda/\Lambda(\nu)) \geq 1$.

Note that $\mu_{j_0}(\Lambda(\nu)) < \mu_{j_0+1}(\Lambda(\nu))$ if $j_0 \geq 1$. Therefore, $\Lambda(\nu)$ appears in the Harder–Narasimhan filtration of Λ as considered by Stuhler [1976] and Grayson [1984], if we include $\{0\}$ as a member of the filtration. In particular, $\Lambda(\nu)$ is the unique lattice in Λ of rank $\mathrm{rk}(\Lambda(\nu))$ and minimal determinant.

Here are two simple properties:

First, for the Euclidean norm $|\cdot|_2$ we claim

$$\log |v|_2 \geq \nu^{\mathrm{rk}(\Lambda/\Lambda(\nu))} \log \det(\Lambda) \quad \text{for all } v \in \Lambda \backslash \Lambda(\nu). \tag{5-2}$$

Indeed, the lattice Λ' generated by $\Lambda(\nu)$ and v contains $\Lambda(\nu)$ strictly. We must have $\mathrm{rk}(\Lambda') > \mathrm{rk}(\Lambda(\nu))$, as $\det(\Lambda')$ would otherwise be strictly less than $\det \Lambda(\nu)$. (This shows in particular that $\Lambda/\Lambda(\nu)$ is torsion free; a well-known property of the Harder–Narasimhan filtration.) So $\mathrm{rk}(\Lambda') = \mathrm{rk}(\Lambda) + 1$ and by convexity of P we find $\log \det(\Lambda') \geq \log \det(\Lambda(\nu)) + \mu_{j_0+1}(\Lambda)$. On the other hand, $\det(\Lambda') \leq \det(\Lambda') \det(\Lambda(\nu) \cap v\mathbb{Z}) \leq \det(\Lambda(\nu)) \det(v\mathbb{Z})$ is well-known, for a proof see [Stuhler 1976, Proposition 2]. We conclude $\log \det(v\mathbb{Z}) \geq \mu_{j_0+1}(\Lambda)$. Now $\det(v\mathbb{Z}) = |v|_2$, so (5-2) follows from (5-1).

Second, (5-1) and $\nu \in \left[0, \frac{1}{2}\right)$ imply

$$\log \det(\Lambda(\nu)) \leq \mu_1(\Lambda) + \cdots + \mu_{j_0}(\Lambda) \leq 2\nu^{1+\mathrm{rk}(\Lambda/\Lambda(\nu))} \log \det(\Lambda). \tag{5-3}$$

We now make things more concrete. Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have order N and set

$$\Lambda_{\boldsymbol{\zeta}} = \{u \in \mathbb{Z}^d : \boldsymbol{\zeta}^u = 1\}. \tag{5-4}$$

We consider the homomorphism $\mathbb{Z}^d \to \mathbb{G}_m$ defined by $u \mapsto \boldsymbol{\zeta}^u$ and see that $\mathbb{Z}^d/\Lambda_{\boldsymbol{\zeta}}$ is isomorphic to the finite subgroup of $\mathbb{G}_m$ generated by the coordinates of $\boldsymbol{\zeta}$. So $\mathbb{Z}^d/\Lambda_{\boldsymbol{\zeta}}$ is cyclic of order N. In particular, $\Lambda_{\boldsymbol{\zeta}}$ is a lattice in $\mathbb{R}^d$ of rank d with $\det(\Lambda_{\boldsymbol{\zeta}}) = [\mathbb{Z}^d : \Lambda_{\boldsymbol{\zeta}}] = N \geq 1$. The saturation

$$\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu) = \{u \in \mathbb{Z}^d : \text{there is } n \in \mathbb{Z}\backslash\{0\} \text{ such that } nu \in \Lambda_{\boldsymbol{\zeta}}(\nu)\} \tag{5-5}$$

of $\Lambda_{\boldsymbol{\zeta}}$ in $\mathbb{Z}^d$ will also be useful for us. It is a lattice of the same rank as $\Lambda_{\boldsymbol{\zeta}}(\nu)$.

For any lattice $\Lambda \subset \mathbb{R}^d$ of positive rank, we set

$$\lambda_1(\Lambda) = \min\{|u| : u \in \Lambda\backslash\{0\}\} \tag{5-6}$$

where as usual $|\cdot|$ denotes the maximum-norm. It is convenient to define $\lambda_1(\{0\}) = \infty$.

Proposition 5.1. *Let $\nu \in \left(0, \frac{1}{4}\right]$ and let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ be of order N. There exists $V \in \mathrm{GL}_d(\mathbb{Z})$ and a decomposition $\boldsymbol{\zeta} = \boldsymbol{\eta}\boldsymbol{\xi}$ with $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ in $\mathbb{G}_m^d$ of finite order E and M, respectively, such that the following holds. We abbreviate $r = \mathrm{rk}(\Lambda_{\boldsymbol{\zeta}}/\Lambda_{\boldsymbol{\zeta}}(\nu)) \in \{1, \dots, d\}$:*

(i) *We have $E \mid N$, $M \mid N$, and $E \le N^{2\nu^{1+r}}$. In particular, $\mathbb{Q}(\boldsymbol{\eta}, \boldsymbol{\xi}) = \mathbb{Q}(\boldsymbol{\zeta})$.*

(ii) *We have $|V| \ll_d N^{2\nu^{1+r}}$ with $\boldsymbol{\xi}^V = (1, \dots, 1, \boldsymbol{\xi}')$ and $\boldsymbol{\xi}' \in \mathbb{G}_m^r$.*

(iii) *If G is a subgroup of $(\mathbb{Z}/M\mathbb{Z})^\times$ there exist $a \in \mathbb{Z}^r$ and $\sigma \in G$ such that $\boldsymbol{\xi}' = \boldsymbol{e}(a\sigma/M)$,*

$$|a| < M \quad \text{and} \quad \frac{|a|}{M} \ll_d \frac{[(\mathbb{Z}/M\mathbb{Z})^\times : G] f_G^{1/2}}{N^{\nu^r/(6d)}}. \tag{5-7}$$

(iv) *With the definition (1-1) we have $\delta(\boldsymbol{\xi}) \ge d^{-1/2} \min\{\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu)), N^{\nu^d/2}\}$.*

Moreover, if $r = d$, or equivalently $\Lambda(\nu) = \{0\}$, then V is the identity matrix.

Proof. We abbreviate $\Lambda = \Lambda_{\boldsymbol{\zeta}}$ as well as $\Lambda(\nu) = \Lambda_{\boldsymbol{\zeta}}(\nu)$ and $\widetilde{\Lambda}(\nu) = \widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu)$. Note $\det(\Lambda) = N$.

We can find a collection of $d - r = \mathrm{rk}(\widetilde{\Lambda}(\nu))$ linearly independent vectors in $\widetilde{\Lambda}(\nu)$ whose norms are at most $\ll_d \det(\widetilde{\Lambda}(\nu))$ by applying Minkowski's second theorem, see Theorem V in Chapter VIII of [Cassels 1959], and using $\lambda_1(\Lambda) \ge 1$. By appending suitable standard basis vectors of $\mathbb{Z}^d$ we find d linearly independent vectors in $\mathbb{Z}^d$. By Corollary 2, Chapter I.2 of [loc. cit.] applied to $\mathbb{Z}^d$ and these vectors we get a basis of $\mathbb{Z}^d$ whose entries have norm $\ll_d \det(\widetilde{\Lambda}(\nu))$. By the said corollary, the original linearly independent vectors can be expressed via an triangular matrix in terms of the new basis vectors. So the first $\mathrm{rk}(\widetilde{\Lambda}(\nu))$ entries of this basis are a basis of the saturated group $\widetilde{\Lambda}(\nu)$. Thus there exists $V \in \mathrm{GL}_d(\mathbb{Z})$ whose first $\mathrm{rk}(\widetilde{\Lambda}(\nu))$ columns constitute a basis of $\widetilde{\Lambda}(\nu)$ and

$$|V| \ll_d \det(\widetilde{\Lambda}(\nu)). \tag{5-8}$$

As $\det(\widetilde{\Lambda}(\nu)) \le \det(\Lambda(\nu))$, the bound for $|V|$ in (ii) follows from (5-3).

We write $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}', \boldsymbol{\xi}')$ where $\boldsymbol{\eta}' \in \mathbb{G}_m^{d-r}$ and $\boldsymbol{\xi}' \in \mathbb{G}_m^r$ both have finite order dividing N. We take $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ from the assertion to equal $(\boldsymbol{\eta}', 1, \dots, 1)^{V^{-1}}$ and $(1, \dots, 1, \boldsymbol{\xi}')^{V^{-1}}$, respectively. So $\boldsymbol{\zeta} = \boldsymbol{\eta}\boldsymbol{\xi}$.

Observe that $[\widetilde{\Lambda}(\nu) : \Lambda(\nu)]\widetilde{\Lambda}(\nu) \subset \Lambda(\nu) \subset \Lambda$. So the first $\mathrm{rk}(\widetilde{\Lambda}(\nu))$ entries of $\boldsymbol{\zeta}^{[\widetilde{\Lambda}(\nu):\Lambda(\nu)]V}$ are $\boldsymbol{\eta}'^{[\widetilde{\Lambda}(\nu):\Lambda(\nu)]} = 1$. This implies that $E = \mathrm{ord}(\boldsymbol{\eta})$ from the assertion satisfies $E \mid [\widetilde{\Lambda}(\nu) : \Lambda(\nu)]$ and thus $E \le \det(\Lambda(\nu)) \le N^{2\nu^{1+r}}$ by (5-3).

To verify (iii) let us fix $v \in \mathbb{Z}^r \setminus \{0\}$ such that $\boldsymbol{\xi}'^v = 1$ and $|v| = \delta(\boldsymbol{\xi}')$. Then $\boldsymbol{\xi}^{V'v} = 1$ where $V' \in \mathrm{Mat}_{dr}(\mathbb{Z})$ consists of the final r columns of V. Raising to the E-th power to kill $\boldsymbol{\eta}$ yields $\boldsymbol{\zeta}^{EV'v} = 1$. Therefore, $EV'v \in \Lambda$. Note that $EV'v \notin \Lambda(\nu)$, indeed otherwise $V'v$ would lie in the saturation $\widetilde{\Lambda}(\nu)$. This is impossible as no nontrivial linear combination of columns of V' lies in $\widetilde{\Lambda}(\nu)$ which is generated by the first $\mathrm{rk}(\widetilde{\Lambda}(\nu))$ columns of V. Thus (5-2) implies $|EV'v|_2 \ge N^{\nu^r}$. By (5-8) we have

$$|EV'v| \ll_d E|V'||v| \ll_d E|V||v| \ll_d [\widetilde{\Lambda}(\nu) : \Lambda(\nu)] \det(\widetilde{\Lambda}(\nu))|v| = \det(\Lambda(\nu))|v|$$

we conclude $N^{\nu^r} \ll_d \det(\Lambda(\nu))|v|$. The determinant bound in (5-3) gives

$$\delta(\boldsymbol{\xi}') = |v| \gg_d N^{\nu^r - 2\nu^{1+r}} \gg_d N^{\nu^r/2}$$

and the last inequality used $\nu \le \frac{1}{4}$.

To complete the proof of (iii) let G be a subgroup of $(\mathbb{Z}/M\mathbb{Z})^\times$ where $M = \mathrm{ord}(\boldsymbol{\xi}) = \mathrm{ord}(\boldsymbol{\xi}')$. By Lemma 3.7 applied to $\boldsymbol{\xi}'$ there are $\sigma \in G$ and $a \in \mathbb{Z}^r$ with $\boldsymbol{\xi}' = \boldsymbol{e}(a\sigma/M)$, $|a| < M$, and

$$\frac{|a|}{M} \ll_d \frac{[(\mathbb{Z}/M\mathbb{Z})^\times : G]^{1/r} f_G^{1/(2r)}}{\delta(\boldsymbol{\xi}')^{1/(3r)}} \ll_d \frac{[(\mathbb{Z}/M\mathbb{Z})^\times : G] f_G^{1/2}}{N^{\nu^r/(6d)}}.$$

It remains to check (iv). Say $v \in \mathbb{Z}^d \backslash \{0\}$ with $\boldsymbol{\xi}^v = 1$ and $|v| = \delta(\boldsymbol{\xi})$. Then $\boldsymbol{\zeta}^v = \boldsymbol{\eta}^v \boldsymbol{\xi}^v = \boldsymbol{\eta}^v$. Thus $Ev \in \Lambda$ and there are two cases to consider. If $v \in \widetilde{\Lambda}(\nu)$, then $\sqrt{d}|v| \geq |v|_2 \geq \lambda_1(\widetilde{\Lambda}(\nu))$ by definition. Otherwise, $v \notin \widetilde{\Lambda}(\nu)$ in which case $Ev \notin \Lambda(\nu)$ by saturation. Here we can use (5-2) and the bound for E from (i) to conclude $|v|_2 \geq E^{-1} N^{\nu^r} \geq N^{\nu^r - 2\nu^{1+r}} \geq N^{\nu^r/2}$. So $|v| \geq |v|_2/\sqrt{d} \geq N^{\nu^r/2}/\sqrt{d} \geq N^{\nu^d/2}/\sqrt{d}$, as claimed in (iv). □

The situation simplifies in the following two cases. If $r = d$, then $\boldsymbol{\xi} = \boldsymbol{\zeta}$, $\boldsymbol{\eta} = 1$, $M = N$, $E = 1$, and V is the identity matrix. If N is a prime, then $E = 1$ as $E \mid N$ and $E \leq N^{2\nu^{1+r}} < N$ by part (i) above. Thus again $\boldsymbol{\xi} = \boldsymbol{\zeta}$ and $\boldsymbol{\eta} = 1$.

6. A preliminary result

Let $d \geq 1$ be an integer.

Definition 6.1. We use the convention $\inf \varnothing = \infty$. For $u \in \mathbb{Z}^d$ we define

$$\rho(u) = \inf\{|v| : v \in \mathbb{Z}^d \backslash \{0\} \text{ and } \langle u, v\rangle = 0\}, \tag{6-1}$$

as usual $|\cdot|$ is the maximum-norm on $\mathbb{R}^d$. Let $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$ be a Laurent polynomial. If the equation $P(\boldsymbol{\eta} z^u) = 0$ has no solution in $\boldsymbol{\eta} \in (\mu_\infty)^d$, $z \in S^1 \backslash \mu_\infty$, and $u \in \mathbb{Z}^d$, we set $\mathcal{B}(P) = 1$. Else wise we set

$$\mathcal{B}(P) = \inf\{B \in \mathbb{N} : \text{ if } \boldsymbol{\eta} \in (\mu_\infty)^d, z \in S^1 \backslash \mu_\infty \text{ is algebraic, and } u \in \mathbb{Z}^d \text{ with } P(\boldsymbol{\eta} z^u) = 0 \text{ then } \rho(u) \leq B\}.$$

Let us spell this out for $d = 1$. Then $\rho(u) = 1$ for $u = 0$ and $\rho(u) = \infty$ otherwise. If P vanishes at a point S^1 of infinite order, then $\mathcal{B}(P) = \infty$. Conversely, if P does not vanish at any point of $S^1 \backslash \mu_\infty$ then we have $\mathcal{B}(P) = 1$. In particular, if $d = 1$ and P is essentially atoral, then $\mathcal{B}(P) = 1$.

Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have order N and say $\nu \in \left(0, \frac{1}{4}\right]$. Below we make use of the canonically determined lattice $\Lambda_{\boldsymbol{\zeta}}(\nu)$ attached to $(\boldsymbol{\zeta}, \nu)$ as in Section 5. Recall that $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu))$ is the least positive Euclidean norm of a vector in the saturation of $\Lambda_{\boldsymbol{\zeta}}(\nu)$ in $\mathbb{Z}^d$. For technical reasons we work with

$$\tilde{\lambda}(\boldsymbol{\zeta}; \nu) = \min\{\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu)), N^{\nu^d/2}\}. \tag{6-2}$$

For example, if $\Lambda_{\boldsymbol{\zeta}}(\nu)$ is $\{0\}$, then the minimum equals $N^{\nu^d/2}$.

An important goal is to generalize Proposition 4.5 to multivariate polynomials. Proposition 6.2 below is a step in this direction.

Proposition 6.2. *Let $K \subset \mathbb{C}$ be a number field, $0 < \nu \le 1/(128d^2)$, and suppose $P \in K[X_1, \dots, X_d]\backslash\{0\}$ has at most k nonzero terms for an integer $k \ge 2$ and satisfies $\mathcal{B}(P) < \infty$. Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have order N and suppose G is a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ with $P(\boldsymbol{\zeta}^\sigma) \neq 0$ for all $\sigma \in G$. Then the following properties hold true with $r = d - \mathrm{rk}(\Lambda_{\boldsymbol{\zeta}}(\nu)) \ge 1$:*

(i) *If $d = 1$, then*

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^\sigma)| = m(P) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^2(1+h(P))}{N^{\nu^r/(20d)}}\right).$$

(ii) *If $d \ge 2$ and $\tilde{\lambda}(\boldsymbol{\zeta};\nu) > d^{1/2}\max\{\mathcal{B}(P), \deg P\}$, then*

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^\sigma)|$$
$$= m(P) + O_{d,k,\nu}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^2(1+h(P))}{N^{\nu^r/(20d)}} + \frac{\deg(P)^{16d^2}}{\tilde{\lambda}(\boldsymbol{\zeta};\nu)^{1/(16(k-1))}}\right). \quad (6\text{-}3)$$

Proof. We may assume that P is nonconstant. Part (i) follows with ample margin from Proposition 4.5 with $Q = P$ and $F = K$. Indeed, we require the standard estimate $d_0(N) \ll_\epsilon N^\epsilon$ which holds for all $\epsilon > 0$. We refrain from stating better bounds in (i) for the purpose of better comparability with the bounds in part (ii).

We split the proof of part (ii) up into 5 steps:

Step 1: Reduction to the univariate case. We write L for the fixed field of G in $\mathbb{Q}(\boldsymbol{\zeta})$. Note that G is the Galois group of $\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/L) = \mathrm{Gal}(L(\boldsymbol{\zeta})/L)$.

By Proposition 5.1 applied to $\boldsymbol{\zeta}$ we obtain $V \in \mathrm{GL}_d(\mathbb{Z})$ and a decomposition $\boldsymbol{\zeta} = \boldsymbol{\eta}\boldsymbol{\xi}$. Let $E = \mathrm{ord}(\boldsymbol{\eta})$ and $M = \mathrm{ord}(\boldsymbol{\xi})$. By (i) of Proposition 5.1 we find

$$E \le N^{2\nu^{1+r}} \quad \text{and thus} \quad M \ge N/E \ge N^{1-2\nu^{1+r}}. \quad (6\text{-}4)$$

The group used in Proposition 5.1(iii) is obtained as follows; we denote it with H to avoid a clash of notation with G from above. Let H be the subgroup of $(\mathbb{Z}/M\mathbb{Z})^\times$ corresponding to $\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\xi})/\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}))$. By Galois theory, see for example [Lang 2002, Theorem VI.1.12], the restriction homomorphism $\mathrm{Gal}(L(\boldsymbol{\xi})/L(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta})) \to \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\xi})/\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}))$ is an isomorphism. Using this isomorphism we will identify H with $\mathrm{Gal}(L(\boldsymbol{\xi})/L(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}))$.

For future reference we estimate the conductor of $H \subset (\mathbb{Z}/M\mathbb{Z})^\times$. The fixed field of H in $\mathbb{Q}(\boldsymbol{\xi})$ is $\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta})$. By the characterization of f_G, the field L is contained in $\mathbb{Q}(\boldsymbol{e}(1/f_G))$. So $\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}) \subset \mathbb{Q}(\boldsymbol{e}(1/M)) \cap \mathbb{Q}(\boldsymbol{e}(1/f_G), \boldsymbol{e}(1/E))$ since $\boldsymbol{\xi}$ has order M and $\boldsymbol{\eta}$ has order E. This final intersection is generated by a root of unity of order $\gcd(M, \mathrm{lcm}(f_G, E))$. We conclude

$$f_H \le \mathrm{lcm}(f_G, E) \le f_G E \le f_G N^{2\nu^{1+r}} \quad (6\text{-}5)$$

having used (6-4).

We use basic Galois theory to compute

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^{\sigma})| = \frac{1}{[L(\boldsymbol{\eta}):L]}\sum_{\tau\in\mathrm{Gal}(L(\boldsymbol{\eta})/L)}\ \sum_{\substack{\sigma\in\mathrm{Gal}(L(\boldsymbol{\xi})/L)\\ \tau|_{L(\boldsymbol{\eta})\cap L(\boldsymbol{\xi})}=\sigma|_{L(\boldsymbol{\eta})\cap L(\boldsymbol{\xi})}}} \frac{1}{\#H}\log|P(\boldsymbol{\eta}^{\tau}\boldsymbol{\xi}^{\sigma})|.$$

Observe that the inner sum is over a coset of $\tilde{\tau}H$ of H inside $(\mathbb{Z}/M\mathbb{Z})^{\times}$; here $\tilde{\tau}\in(\mathbb{Z}/M\mathbb{Z})^{\times}$ restricts to the restriction $\tau|_{L(\boldsymbol{\eta})\cap L(\boldsymbol{\xi})}$. Below, τ is as in the outer sum. The inner sum equals

$$S_{\tau} = \frac{1}{\#H}\sum_{\sigma\in\tilde{\tau}H}\log|P(\boldsymbol{\eta}^{\tau}\boldsymbol{\xi}^{\sigma})| = \frac{1}{\#H}\sum_{\sigma\in H}\log|P(\boldsymbol{\eta}^{\tau}\boldsymbol{\xi}^{\tilde{\tau}\sigma})|, \tag{6-6}$$

and the complete sum is

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^{\sigma})| = \frac{1}{[L(\boldsymbol{\eta}):L]}\sum_{\tau\in\mathrm{Gal}(L(\boldsymbol{\eta})/L)} S_{\tau}. \tag{6-7}$$

By Proposition 5.1(iii) applied to H we get $a\in\mathbb{Z}^r$ satisfying (5-7) and $\sigma_0\in H$ with

$$\boldsymbol{\xi}^{V} = (1,\dots,1,\boldsymbol{e}(a\sigma_0/M)).$$

We extend a to the left by $d-r$ zeros and obtain a row vector $(0,a)\in\mathbb{Z}^d$. We set $u=(0,a)V^{-1}\in\mathbb{Z}^d$ and use Proposition 5.1(ii) to get $\boldsymbol{\xi}=\boldsymbol{e}(u\sigma_0/M)$. Let us set

$$Q = P(\boldsymbol{\eta}^{\tau}X^{u})X^{l} \tag{6-8}$$

in the unknown X; it depends on τ and the exponent l is chosen to make sure that Q is a polynomial. So $0\neq|P(\boldsymbol{\eta}^{\tau}\boldsymbol{\xi}^{\tilde{\tau}\sigma})| = |Q(\boldsymbol{e}(\tilde{\tau}\sigma\sigma_0/M))|$ and in particular $Q\neq 0$. We may assume that $Q(0)\neq 0$. The coefficients of Q lie in $F=K(\boldsymbol{\eta})$ and Q has at most k nonzero terms as P has at most this many nonzero terms. All this allows us to rewrite (6-6) using a univariate polynomial, σ_0 above is absorbed by the sum

$$S_{\tau} = \frac{1}{\#H}\sum_{\sigma\in H}\log|Q(\boldsymbol{e}(\tilde{\tau}\sigma/M))|. \tag{6-9}$$

Step 2: Nonvanishing of Q on $S^1\backslash\mu_\infty$. Suppose $w\in\mathbb{Z}^d\backslash\{0\}$ satisfies $\langle u,w\rangle=0$ and $|w|=\rho(u)$. Recall that $\boldsymbol{\xi}=\boldsymbol{e}(u\sigma_0/M)$, so $\boldsymbol{\xi}^{w}=1$. Thus $|w|\geq\delta(\boldsymbol{\xi})$ and Proposition 5.1(iv) together with (6-2) yield

$$\rho(u)=|w|\geq d^{-1/2}\tilde{\lambda}(\boldsymbol{\zeta};\nu). \tag{6-10}$$

Let $z\in S^1\backslash\mu_\infty$ be algebraic. If $Q(z)=0$ then $P(\boldsymbol{\eta}^{\tau}z^{u})=0$ by (6-8). By Definition 6.1 we have $\rho(u)\leq\mathcal{B}(P)$. This and (6-10) contradict the lower bound $\tilde{\lambda}(\boldsymbol{\zeta};\nu)>d^{1/2}\mathcal{B}(P)$ in the hypothesis. Hence $Q(z)\neq 0$.

Thus Q, having algebraic coefficients, does not vanish at any point of $S^1\backslash\mu_\infty$. As $\rho(u)>1$ we also have $u\neq 0$.

Step 3: Bounding quantities in preparation for Proposition 4.5. This step is mainly bookkeeping. We aim to apply Proposition 4.5 to Q, the root of unity $\boldsymbol{e}(\tilde{\tau}/M)$, and the subgroup $H \subset (\mathbb{Z}/M\mathbb{Z})^\times$ to determine the asymptotic behavior of S_τ. To proceed we bound the various quantities below separately:

$$\begin{aligned}
[(\mathbb{Z}/M\mathbb{Z})^\times : H] &\le [(\mathbb{Z}/N\mathbb{Z})^\times : G] N^{2\nu^{1+r}},\\
f_H &\le f_G N^{2\nu^{1+r}},\\
\deg(Q) &\ll_d \deg(P) \min\{[(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} N^{1-\nu^r/(10d)}, N^2\},\\
h(Q) &= h(P),\\
[K(\boldsymbol{\eta}) : \mathbb{Q}] = [F : \mathbb{Q}] &\le [K : \mathbb{Q}] N^{2\nu^{1+r}}.
\end{aligned} \tag{6-11}$$

Note that $\#H = [\mathbb{Q}(\boldsymbol{\xi}) : \mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta})] = [\mathbb{Q}(\boldsymbol{\xi}) : \mathbb{Q}]/[\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}) : \mathbb{Q}] \ge [\mathbb{Q}(\boldsymbol{\xi}) : \mathbb{Q}]/[L(\boldsymbol{\eta}) : \mathbb{Q}]$ and since $[\mathbb{Q}(\boldsymbol{\xi}) : \mathbb{Q}] = \#(\mathbb{Z}/M\mathbb{Z})^\times$ we find $[(\mathbb{Z}/M\mathbb{Z})^\times : H] \le [L(\boldsymbol{\eta}) : \mathbb{Q}] \le [L : \mathbb{Q}]E$. The first bound follows from (6-4) and as $[L : \mathbb{Q}] = [(\mathbb{Z}/N\mathbb{Z})^\times : G]$.

We already proved the bound for f_H in (6-5).

Next comes $\deg(Q)$. Observe that

$$\begin{aligned}
\deg(Q) &\ll_d |a||V^{-1}| \deg(P) \ll_d |a||V|^{d-1} \deg(P)\\
&\ll_d [(\mathbb{Z}/M\mathbb{Z})^\times : H] f_H^{1/2} \deg(P) N^{1+2(d-1)\nu^{1+r}-\nu^r/(6d)}\\
&\ll_d [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \deg(P) N^{1+2\nu^{1+r}+\nu^{1+r}+2(d-1)\nu^{1+r}-\nu^r/(6r)}
\end{aligned}$$

having used the bounds in Proposition 5.1, $M \le N$, and the first two bounds in (6-11). As $\nu \le 1/(128d^2)$ the exponent of N is at most $1 + (2d+1)\nu^{1+r} - \nu^r/(6d) \le 1 - \nu^r/(10d)$ and thus we obtain

$$\deg(Q) \ll_d [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \deg(P) N^{1-\nu^r/(10d)}$$

which is part of the third inequality in (6-11). The bound $\deg(Q) \ll_d \deg(P)N^2$ is proved similarly, but requires only the trivial estimate $|a| < M \le N$ from (5-7) and $|V^{-1}| \ll_d N^{2d\nu^{1+r}}$.

We claim that the coefficients of $P(\boldsymbol{\eta}^\tau X^u)$ are equal to the coefficients of P up to multiplication by a root of unity. In view of the definition of the height (2-4) this will imply the fourth claim in (6-11). Indeed, it suffices to rule out that two distinct monomials in P lead to the same power of X after the substitution. Hence it suffices to verify $\rho(u) > \deg P$. But this follows from (6-10) and as $\tilde{\lambda}(\boldsymbol{\zeta}; \nu) > d^{1/2} \deg P$ by hypothesis.

The degree of the number field F containing the coefficients of Q satisfies

$$[F : \mathbb{Q}] = [K(\boldsymbol{\eta}) : \mathbb{Q}] \le [K : \mathbb{Q}][\mathbb{Q}(\boldsymbol{\eta}) : \mathbb{Q}] \le [K : \mathbb{Q}]E \le [K : \mathbb{Q}] N^{2\nu^{1+r}}$$

where we used (6-4). This implies the fifth claim in (6-11).

Step 4: Applying Proposition 4.5 in the univariate case. Our aim is to determine the asymptotics of (6-9). We use the bounds from the last step to control the error term in (4-8) arise in Proposition 4.5

applied to F, Q, $\boldsymbol{e}(\tilde{\tau}/M)$ of order M, and to H. By (6-11) the error is

$$\ll [F:\mathbb{Q}]^2[(\mathbb{Z}/M\mathbb{Z})^\times : H] f_H^{1/2} \deg(Q)(\log(2\deg Q) + h(Q)) \frac{(\log 2M)^3 d_0(M)}{M}$$

$$\ll_d [K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)(\log(2N^2 \deg P) + h(P)) N^{9\nu^{1+r}+1-\nu^r/(10d)} \frac{(\log 2M)^3 d_0(M)}{N}$$

where we use $\deg Q \ll N^2 \deg P$ to bound $\log(2\deg Q)$ from above and the lower bound for M in (6-4).

The exponent of N is $9\nu^{1+r} - \nu^r/(10d) \le -\nu^r/(19d)$ as $\nu \le 1/(128d^2) \le 1/(256d)$. As $M \mid N$ we find $d_0(M) \le d_0(N)$. As in the proof of (i) we use $d_0(N) \ll_\epsilon N^\epsilon$ for all ϵ. We also anticipate $\log(2N^2)$ coming from $\log(2N^2 \deg P)$ to find

$$\log(2N^2) N^{9\nu^{1+r}-\nu^r/(10d)} (\log 2M)^3 d_0(M) \ll_{d,\nu} N^{-\nu^r/(20d)}.$$

Using the crude inequality $\log \deg P \le \deg P$ the error term is thus

$$\ll_{d,\nu} [K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^2(1+h(P)) N^{-\nu^r/(20d)}.$$

Applying Proposition 4.5 and recalling $m(Q) = m(P(\boldsymbol{\eta}^\tau X^u))$ we find

$$S_\tau = m(P(\boldsymbol{\eta}^\tau X^u)) + O_{d,\nu}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^2(1+h(P))}{N^{\nu^r/(20d)}} \right). \tag{6-12}$$

Step 5: Applying a quantitative version of Lawton's theorem. To determine the asymptotics of the Mahler measure we apply our quantitative variant of Lawton's theorem, Theorem A.1 to $P(\boldsymbol{\eta}^\tau(X_1,\dots,X_d)) \neq 0$ and u. This polynomial has the same degree and number of terms as P. The exponent vector satisfies $\rho(u) \ge d^{-1/2}\tilde{\lambda}(\boldsymbol{\zeta};\nu)$ by (6-10). Our hypothesis implies $\rho(u) > \deg P$, as required by Theorem A.1. We find

$$m(P(\boldsymbol{\eta}^\tau X^u)) = m(P(\boldsymbol{\eta}^\tau(X_1,\dots,X_d))) + O_{d,k}\left(\frac{\deg(P)^{16d^2}}{\tilde{\lambda}(\boldsymbol{\zeta};\nu)^{1/(16(k-1))}} \right). \tag{6-13}$$

The Mahler measure of P and $P(\boldsymbol{\eta}^\tau(X_1,\dots,X_d))$ are equal as translating by $\boldsymbol{\eta}^\tau \in (S^1)^d$ does not affect the value of the integral (1-4).

By combining (6-12) and (6-13) we conclude

$$S_\tau = m(P) + O_{d,k,\nu}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^2(1+h(P))}{N^{\nu^r/(20d)}} + \frac{\deg(P)^{16d^2}}{\tilde{\lambda}(\boldsymbol{\zeta};\nu)^{1/(16(k-1))}} \right).$$

Part (ii) of the proposition follows from (6-7). □

We now explain why the situation simplifies when the order N of $\boldsymbol{\zeta}$ is a prime number. In this case, after the proof of Proposition 5.1 we observed that $\boldsymbol{\eta} = 1$ and $\boldsymbol{\zeta} = \boldsymbol{\xi}$. In the proof above, inequality (6-10) can be replaced by $\rho(u) \ge \delta(\boldsymbol{\zeta})$. So the hypothesis on $\boldsymbol{\zeta}$ in (ii) of the proposition can be replaced by $\delta(\boldsymbol{\zeta}) > \max\{\mathcal{B}(P), \deg P\}$; see also the argument near (6-13). This is certainly satisfied for $\delta(\boldsymbol{\zeta}) \to \infty$. Moreover, $\tilde{\lambda}(\boldsymbol{\zeta};\nu)$ can be replaced by $\delta(\boldsymbol{\zeta})$ in (6-3). From this point it is not difficult to deduce Theorem 1.1 when N is a prime.

The remaining argument is required to treat general N. We need to keep track of extra information such as $[K:\mathbb{Q}]$, $[(\mathbb{Z}/N\mathbb{Z})^\times : G]$, f_G, and the dependency on P to anticipate a monomial change of coordinates.

7. Equidistribution

Proposition 6.2 closes in on Theorem 1.1. Indeed, suppose that for some choice of ν the value $\tilde{\lambda}(\boldsymbol{\zeta};\nu)$ grows polynomially in $\delta(\boldsymbol{\zeta})$. Then the error term of (6-3) tends to 0 as $\delta(\boldsymbol{\zeta})\to\infty$ and we are done.

However, consider the following example, already found in the beginning of Section 5. Suppose $n\ge 2$ and $\boldsymbol{\zeta}_p$ and $\boldsymbol{\zeta}_{p^n}$ are roots of unity of order p and p^n, respectively. Say $\boldsymbol{\zeta}=(\boldsymbol{\zeta}_p,\boldsymbol{\zeta}_{p^n})$, it has order p^n. The lattice $\Lambda_{\boldsymbol{\zeta}}$ contains $(p,0)^t$ and this vector has minimal positive Euclidean norm in $\Lambda_{\boldsymbol{\zeta}}$. For n large enough in terms of ν we have $\Lambda(\nu)=(p,0)^t\mathbb{Z}$ and $\widetilde{\Lambda}(\nu)=(1,0)^t\mathbb{Z}$. Thus $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu))=1$ and this yields $\tilde{\lambda}(\boldsymbol{\zeta};\nu)=1$.

This example suggests a monomial change of coordinates which we will do in the next section. In the current section we lay the groundwork for this change of coordinates.

7A. ***Numerical integration.*** We require a higher dimensional replacement of the Koksma bound [Harman 1998, Theorem 5.4]. The classical analog is called the Koksma–Hlawka Inequality and applies to functions of bounded variation in the sense of Hardy and Krause. Let $\theta: U\to\mathbb{R}$ be a function whose domain U is a nonempty subset of $\mathbb{R}^d$. In this subsection we use the *modulus of continuity* of θ defined by

$$\omega(\theta;t)=\sup_{\substack{x,y\in U\\ |x-y|\le t}}|\theta(x)-\theta(y)| \tag{7-1}$$

for all $t\ge 0$; as usual $|\cdot|$ denotes the maximum-norm on $\mathbb{R}^d$. We define $\omega(\theta;t)=0$ if $U=\varnothing$. We will use it to estimate a mean in terms of the corresponding integral in Proposition 7.1. Hlawka [1971] has a related and more precise result. For the reader's convenience we give a self-contained treatment that suffices for our purposes.

Proposition 7.1. *Let $\theta:[0,1]^d\to\mathbb{R}$ be a continuous function and let $x_1,\dots,x_n\in[0,1)^d$ with discrepancy $\mathcal{D}=\mathcal{D}(x_1,\dots,x_n)$. Then*

$$\left|\frac{1}{n}\sum_{i=1}^{n}\theta(x_i)-\int_{[0,1)^d}\theta(x)\,dx\right|\le(1+2^{d+1})\omega(\theta,\mathcal{D}^{1/(d+1)}). \tag{7-2}$$

Proof. Both sides of (7-2) are invariant under adding a constant function to θ. So we may assume $\theta(0)=0$.

Let $T\ge 1$ be an integral parameter to be determined below. We write $[0,1)^d$ as a disjoint union of T^d half-open hypercubes Q_j with side length $1/T$. Let $\overline{Q_j}$ denote the closure of Q_j in $[0,1]^d$. The mean value theorem tells us that for each j there exists $y_j\in\overline{Q_j}$ such that $\int_{Q_j}\theta(x)\,dx=\operatorname{vol}(Q_j)\theta(y_j)=T^{-d}\theta(y_j)$.

For each j we write $n_j=\#\{i\in\{1,\dots,n\}: x_i\in Q_j\}$. So

$$\frac{1}{n}\left|\sum_{i=1}^{n}\theta(x_i)-\sum_j n_j\theta(y_j)\right|\le\frac{1}{n}\sum_j\sum_{\substack{i=1\\ x_i\in Q_j}}^{n}|\theta(x_i)-\theta(y_j)|\le\frac{1}{n}\sum_j\omega(\theta;1/T)n_j=\omega(\theta;1/T). \tag{7-3}$$

On the other hand, $\frac{1}{n}\sum_j n_j\theta(y_j)$ equals

$$\sum_j \frac{n_j}{n}T^d\int_{Q_j}\theta(x)\,dx=\sum_j(1+\delta_jT^d)\int_{Q_j}\theta(x)\,dx=\int_{[0,1)^d}\theta(x)\,dx+T^d\sum_j\delta_j\int_{Q_j}\theta(x)\,dx$$

where $\delta_j=n_j/n-T^{-d}$. The definition of discrepancy implies $|\delta_j|\le\mathcal{D}$. Hence

$$\left|\frac{1}{n}\sum_j n_j\theta(y_j)-\int_{[0,1)^d}\theta(x)\,dx\right|\le T^d\mathcal{D}\int_{[0,1)^d}|\theta(x)|\,dx\le T^{d+1}\mathcal{D}\omega(\theta;1/T)\tag{7-4}$$

where we used $|\theta(x)|\le T\omega(\theta;1/T)$ for all $x\in[0,1]^d$; recall that $\theta(0)=0$.

We apply the triangle inequality to (7-3) and (7-4) and conclude that the left-hand side of (7-2) is at most $(1+T^{d+1}\mathcal{D})\omega(\theta;1/T)$. To complete the proof observe that $0<\mathcal{D}\le 1$ and fix $T=\lceil\mathcal{D}^{-1/(d+1)}\rceil$ which satisfies $\mathcal{D}^{-1/(d+1)}\le T\le\mathcal{D}^{-1/(d+1)}+1$. □

7B. ***Averaging the Mahler measure.*** This subsection is purely in the complex setting. Let $P\in\mathbb{C}[X_1,\dots,X_d]\backslash\{0\}$ have at most $k\ge 2$ nonzero terms, where k is an integer.

Let $l\in\{1,\dots,d-1\}$. For $x\in\mathbb{R}^l$ we define $P_{\boldsymbol{e}(x)}=P(\boldsymbol{e}(x),X_1,\dots,X_{d-l})\in\mathbb{C}[X_1,\dots,X_{d-l}]$. Next we construct an auxiliary Laurent polynomial $\widehat{P}$ in l variables whose value at $\boldsymbol{e}(x)$ is comparable to $|P_{\boldsymbol{e}(x)}|$. For $i\in\mathbb{Z}^{d-l}$ we denote $p_i\in\mathbb{C}[X_1,\dots,X_l]$ the coefficients of P, taken as a Laurent polynomial in $X_{l+1},\dots,X_d$, and define

$$\widehat{P}=\sum_i p_i(X_1,\dots,X_l)\overline{p_i}(X_1^{-1},\dots,X_l^{-1})\in\mathbb{C}[X_1^{\pm1},\dots,X_l^{\pm1}]\tag{7-5}$$

where the bar denotes complex conjugation.

Lemma 7.2. *In the notation above the following properties hold true*:

(i) *The Laurent polynomial $\widehat{P}$ has at most k^2 nonzero terms.*

(ii) *The product $(X_1\cdots X_l)^{\deg P}\widehat{P}$ is a polynomial of degree at most $(l+1)\deg P$.*

Proof. If each p_i consists of k_i nonzero terms, then $\widehat{P}$ consists of at most $\sum_i k_i^2$ terms. Since $\sum_i k_i\le k$ we find that $\widehat{P}$ has at most k^2 nonzero terms. This implies part (i).

Part (ii) follows from (7-5). □

Observe that $\widehat{P}(\boldsymbol{e}(x))=\sum_i|p_i(\boldsymbol{e}(x))|^2\ge 0$. As $|P_{\boldsymbol{e}(x)}|$ is the maximum of $|p_i(\boldsymbol{e}(x))|$ as i varies, we find

$$\frac{1}{k^{1/2}}\widehat{P}(\boldsymbol{e}(x))^{1/2}\le|P_{\boldsymbol{e}(x)}|\le\widehat{P}(\boldsymbol{e}(x))^{1/2}.\tag{7-6}$$

So $P_{\boldsymbol{e}(x)}=0$ if and only if $\widehat{P}(\boldsymbol{e}(x))=0$.

The main result of this subsection is:

Proposition 7.3. *Assume $P \in \mathbb{C}[X_1, \dots, X_d]\backslash\mathbb{C}$ has at most k nonzero terms for an integer $k \geq 2$. Let $l \in \{1, \dots, d-1\}$ and let $\widehat{P}$ be as above. Suppose $x_1, \dots, x_n \in [0,1)^l$ with discrepancy $\mathcal{D} = \mathcal{D}(x_1, \dots, x_n)$. If $P_{\boldsymbol{e}(x_i)} \neq 0$ for all $i \in \{1, \dots, n\}$, then*

$$\frac{1}{n}\sum_{i=1}^{n} m(P_{\boldsymbol{e}(x_i)}) = m(P) + O_{d,k}\left(\deg(P)\mathcal{D}^{1/(16(d+1)k^2)} + \left|m(\widehat{P}) - \frac{1}{n}\sum_{i=1}^{n}\log\widehat{P}(\boldsymbol{e}(x_i))\right|\right). \tag{7-7}$$

By a theorem of Boyd [1998], the Mahler measure is a continuous function in the coefficients of a nonzero polynomial of fixed degree (below in Lemma A.5 we prove that it is even Hölder continuous). Therefore, if the $P_{\boldsymbol{e}(x_i)}$ in the proposition above are uniformly bounded away from 0, then the average on the left in (7-7) converges to the integral $\int_{[0,1)^l} m(P_{\boldsymbol{e}(x)})\,dx$ as the discrepancy tends to 0. But even when $|P| = 1$ it is conceivable that $|P_{\boldsymbol{e}(x_i)}|$ is small for some x_i, then $P_{\boldsymbol{e}(x_i)}$ is near the Mahler measure's logarithmic singularity. This happens if and only if $\widehat{P}(\boldsymbol{e}(x_i))$ is small by (7-6). The proposition states that we can handle the mean for arbitrary x_i if we can control the logarithmic mean of $\widehat{P}$ over the $\boldsymbol{e}(x_i)$.

The proof follows a series of lemmas. We first note a useful property of the modulus of continuity as defined in (7-1). Let $\theta : [0,1]^d \to \mathbb{R}\cup\{-\infty\}$ be a function and $c \in \mathbb{R}$, such that $\theta_c(x) = \max\{c, \theta(x)\}$ defines a continuous function $[0,1]^d \to \mathbb{R}$. We claim that

$$\omega(\theta_c; t) \leq \omega(\theta|_{\theta^{-1}((c,\infty))}; t) \quad \text{for all } t \geq 0. \tag{7-8}$$

This inequality follows by definition if $\theta^{-1}((c,\infty))$ is empty. Say $x, y \in [0,1]^d$ with $|x-y| \leq t$. To bound $|\theta_c(x) - \theta_c(y)|$ from above by the right-hand side of (7-8) we may assume $\theta_c(x) > c = \theta_c(y)$. By continuity of θ_c there is for all small enough $\epsilon > 0$ a $z \in [0,1]^d$ on the line segment connecting x and y with $c+\epsilon = \theta_c(z) = \theta(z)$. Then $|\theta_c(x)-\theta_c(y)| = |\theta(x)-c| \leq |\theta(x)-\theta(z)| + |\theta(z)-c| \leq \omega(\theta|_{\theta^{-1}((c,\infty))}; t) + \epsilon$. Our claim (7-8) follows as ϵ can be made arbitrarily small.

Let P and k be as in Proposition 7.3 and assume in addition that $|P| = 1$.

Lemma 7.4. *let $x_1, \dots, x_n \in [0,1)^d$ have discrepancy $\mathcal{D} = \mathcal{D}(x_1, \dots, x_n)$. If $r \in (0,1]$, then*

$$\frac{1}{n}\#\{i \in \{1, \dots, n\} : |P(\boldsymbol{e}(x_i))| \leq r\} \ll_{d,k} r^{1/(2k)} + \deg(P)\mathcal{D}^{1/(d+1)}/r. \tag{7-9}$$

Proof. For $x \in [0,1]^d$ we set

$$\chi(x) = \max\{0, 2 - |P(\boldsymbol{e}(x))|/r\}$$

and this defines a continuous function on $[0,1]^d$ with values in $[0,2]$.

We note that $\chi(x) \geq 1$ if $|P(\boldsymbol{e}(x))| \leq r$. As χ is nonnegative the average $\frac{1}{n}\sum_{i=1}^{n}\chi(x_i)$ is at least the proportion of the i among $\{1, \dots, n\}$ such that $|P(\boldsymbol{e}(x_i))| \leq r$. On the other hand, Lemma A.3(i) implies

$$\int_{[0,1)^d} \chi(x)\,dx \leq 2\mathrm{vol}(\{x \in [0,1)^d : |P(\boldsymbol{e}(x))| < 2r\}) \ll_{d,k} r^{1/(2(k-1))} \ll_{d,k} r^{1/(2k)}. \tag{7-10}$$

We will apply Proposition 7.1 to bound the proportion on the left in (7-9). Say $t > 0$, let us verify

$$\omega(\chi; t) \ll_{d,k} \deg(P)t/r. \tag{7-11}$$

We apply (7-8) to $\theta(x) = 2 - |P(\boldsymbol{e}(x))|/r$ and $c = 0$. Say $x, y \in \theta^{-1}((0,\infty))$ with $|x-y| \le t$, so in particular $|P(\boldsymbol{e}(x))| < 2r$ and $|P(\boldsymbol{e}(y))| < 2r$. Then $|\theta(x) - \theta(y)| = |P(\boldsymbol{e}(x)) - P(\boldsymbol{e}(y))|/r \ll_{d,k} \deg(P)t/r$, where we used $|x-y| \le t$ and $|P| = 1$. We obtain (7-11).

Let us set $t = \mathcal{D}^{1/(d+1)}$. We apply numerical integration, Proposition 7.1, and use (7-10) to conclude the proof. □

In the next lemma we truncate the singularity of $x \mapsto \log|P(\boldsymbol{e}(x))|$ using a parameter r and bound the modulus of continuity of the resulting function.

Lemma 7.5. *Let $r \in (0,1]$, for $x \in [0,1]^d$ we define $\psi(x) = \max\{\log r, \log|P(\boldsymbol{e}(x))|\}$ as above (7-8). Then $\psi : [0,1]^d \to \mathbb{R}$ is continuous and for all $t > 0$ we have*

$$\omega(\psi; t) \ll_{d,k} \frac{\deg(P)t}{r}.$$

Proof. Clearly, ψ is continuous on $[0,1]^d$. We apply (7-8) to $\theta(x) = \log|P(\boldsymbol{e}(x))|$ and $c = \log r$. Say $x, y \in [0,1]^d$ with $|P(\boldsymbol{e}(x))| \ge |P(\boldsymbol{e}(y))| \ge r$ and $|x-y| \le t$. Then as in the proof of Lemma 7.4 we find $\big||P(\boldsymbol{e}(x))/P(\boldsymbol{e}(y))| - 1\big| \ll_{d,k} \deg(P)t/|P(\boldsymbol{e}(y))| \ll_{d,k} \deg(P)t/r$. Applying the logarithm and using $0 \le \log s \le s - 1$ for all $s \ge 1$ yields

$$\big|\log|P(\boldsymbol{e}(x))| - \log|P(\boldsymbol{e}(y))|\big| \ll_{d,k} \frac{\deg(P)t}{r},$$

as desired. □

Lemma 7.6. *We keep the notation of Lemma 7.5. Then*

$$\left| m(P) - \int_{[0,1)^d} \psi(x)\, dx \right| \ll_{d,k} r^{1/(4k)}.$$

Proof. The absolute value in question is

$$\mathcal{E} = \left| \int_{\Sigma} \log|P(\boldsymbol{e}(x))|\, dx - \mathrm{vol}(\Sigma)\log r \right|$$

where $\Sigma = S(P,r) = \{x \in [0,1)^d : |P(\boldsymbol{e}(x))| < r\}$ in the notation of (A-2). Hence $\mathrm{vol}(\Sigma) \ll_{d,k} r^{1/(2(k-1))}$ by Lemma A.3(i). So

$$\mathcal{E} \ll_{d,k} \int_{\Sigma} \big|\log|P(\boldsymbol{e}(x))|\big|\, dx + r^{1/(2k)}$$

as $r \le 1$. To bound the final integral we use Lemma A.4 which implies $\mathcal{E} \ll_{d,k} r^{1/(4(k-1))} + r^{1/(2k)} \ll_{d,k} r^{1/(4k)}$. □

Lemma 7.7. *Let $x_1, \dots, x_n \in [0,1)^d$ with $P(\boldsymbol{e}(x_i)) \ne 0$ for all i and discrepancy $\mathcal{D} = \mathcal{D}(x_1, \dots, x_n)$. We set*

$$\epsilon = \left| m(P) - \frac{1}{n}\sum_{i=1}^{n} \log|P(\boldsymbol{e}(x_i))| \right|.$$

If $r \in (0, 1]$, *then*

$$\frac{1}{n} \sum_{|P(\boldsymbol{e}(x_i))|<r} \big|\log|P(\boldsymbol{e}(x_i))|\big| \ll_{d,k} \deg(P)\mathcal{D}^{1/(d+1)}r^{-2} + r^{1/(4k)} + \epsilon.$$

Proof. By the triangle inequality and with ψ as in Lemma 7.5 we have

$$\left|\frac{1}{n}\sum_{i=1}^{n}\psi(x_i) - \log|P(\boldsymbol{e}(x_i))|\right| \le \left|\frac{1}{n}\sum_{i=1}^{n}\psi(x_i) - \int_{[0,1)^d}\psi(x)\,dx\right| + \left|\int_{[0,1)^d}\psi(x)\,dx - m(P)\right| + \epsilon.$$

We use Proposition 7.1 and Lemma 7.5 with $t = \mathcal{D}^{1/(d+1)}$ to bound the first term on the right by $\ll_{d,k} \deg(P)\mathcal{D}^{1/(d+1)}/r$. The second term is $\ll_{d,k} r^{1/(4k)}$ by Lemma 7.6.

The term on the left equals $\frac{1}{n}\sum_{|P(\boldsymbol{e}(x_i))|<r}(\log r - \log|P(\boldsymbol{e}(x_i))|)$. Observe that $-\log|P(\boldsymbol{e}(x_i))| = \big|\log|P(\boldsymbol{e}(x_i))|\big|$ in this sum as $r \le 1$. We rearrange and find

$$\frac{1}{n} \sum_{|P(\boldsymbol{e}(x_i))|<r} \big|\log|P(\boldsymbol{e}(x_i))|\big| \ll_{d,k} \deg(P)\mathcal{D}^{1/(d+1)}r^{-1} + r^{1/(4k)} + \frac{|\log r|}{n}\Bigg(\sum_{|P(\boldsymbol{e}(x_i))|<r} 1\Bigg) + \epsilon.$$

By Lemma 7.4, the term corresponding to the sum over i on the right is

$$\ll_{d,k} r^{1/(2k)}|\log r| + \deg(P)\mathcal{D}^{1/(d+1)}r^{-1}|\log r|.$$

Combining our bounds and absorbing $|\log r|$ in an appropriate power of r^{-1} we find

$$\frac{1}{n} \sum_{|P(\boldsymbol{e}(x_i))|<r} \big|\log|P(\boldsymbol{e}(x_i))|\big| \ll_{d,k} \deg(P)\mathcal{D}^{1/(d+1)}r^{-2} + r^{1/(4k)} + \epsilon,$$

as desired. □

After this warming-up we prove variants of Lemmas 7.5 and 7.6 where $\log|\cdot|$ is replaced by the Mahler measure. We also truncate at the parameter r.

Lemma 7.8. *Let* $r \in (0, 1]$, *for* $x \in [0, 1]^l$ *we define* $\mu(x) = \max\{\log r, m(P_{\boldsymbol{e}(x)})\}$ *as above (7-8) where we interpret the Mahler measure of* 0 *as* $-\infty$. *Then* $\mu : [0, 1]^l \to \mathbb{R}$ *is continuous and for all* $t > 0$ *we have*

$$\omega(\mu; t) \ll_{d,k} \left(\frac{\deg(P)t}{r}\right)^{1/(8k)}(1 + |\log r|).$$

Proof. By Boyd's theorem [1998] the Mahler measure is continuous on the space of nonzero polynomials of bounded degree. Thus μ is continuous on $[0, 1]^l$. Observe that $\omega(\mu; t) \ll_k 1 + |\log r|$ as $m(P_{\boldsymbol{e}(x)}) \ll_k 1$ by (2-1) and $|P| = 1$. So we may assume that $\deg(P)t/r$ is sufficiently small in terms of d and k.

We again use (7-8), this time with $\theta(x) = m(P_{\boldsymbol{e}(x)})$ and $c = \log r$. Let $x, y \in [0, 1]^d$ with $m(P_{\boldsymbol{e}(x)}) \ge \log r$ and $m(P_{\boldsymbol{e}(y)}) \ge \log r$ and $|x - y| \le t$. Then $|P_{\boldsymbol{e}(x)}| \gg_k r$ and $|P_{\boldsymbol{e}(y)}| \gg_k r$ by (2-1). As in the proof of Lemma 7.4 we find $|P_{\boldsymbol{e}(x)} - P_{\boldsymbol{e}(y)}| \ll_{d,k} \deg(P)t$. Since $\deg(P)t/r$ is smaller than some prescribed constant depending only on d and k we may assume $|P_{\boldsymbol{e}(x)} - P_{\boldsymbol{e}(y)}|/\min\{|P_{\boldsymbol{e}(x)}|, |P_{\boldsymbol{e}(y)}|\} \le \frac{1}{2}$. Lemma A.5

implies

$$|m(P_{\boldsymbol{e}(x)}) - m(P_{\boldsymbol{e}(y)})| \ll_{d,k} \left(\frac{|P_{\boldsymbol{e}(x)} - P_{\boldsymbol{e}(y)}|}{\min\{|P_{\boldsymbol{e}(x)}|, |P_{\boldsymbol{e}(y)}|\}} \right)^{1/(8(k-1))} \ll_{d,k} \left(\frac{\deg(P)t}{r} \right)^{1/(8(k-1))},$$

as desired. □

Before continuing we recall the p_i and the auxiliary Laurent polynomial $\widehat{P}$ determined by P and l in (7-5). By Lemma 7.2(i) $\widehat{P}$ has at most k^2 nonzero terms, so $\sup_{x\in[0,1]^l} |\widehat{P}(\boldsymbol{e}(x))| \le k^2 |\widehat{P}|$. There exists i with $|p_i| = |P| = 1$. The definition of the Mahler measure implies

$$m(p_i) \le \sup_{x\in[0,1]^l} \log|p_i(\boldsymbol{e}(x))| \le \tfrac{1}{2} \sup_{x\in[0,1]^l} \log|\widehat{P}(\boldsymbol{e}(x))| \le \tfrac{1}{2}(2\log k + \log|\widehat{P}|).$$

Using $|p_i| = 1$ and the theorem of Dobrowolski and Smyth, Theorem 2.1, we conclude $m(p_i) \ge -(k-2)\log 2$. Thus $|\widehat{P}| \gg_k 1$. Bounding $|\widehat{P}|$ from above is more straight-forward. Indeed, $|\widehat{P}| \ll_k 1$ by (7-5) and since $|P| = 1$. Therefore,

$$1 \ll_k |\widehat{P}| \ll_k 1. \tag{7-12}$$

We let $\widetilde{P}$ denote the polynomial from Lemma 7.2(ii) divided by $|\widehat{P}|$, so $|\widetilde{P}| = 1$.

Lemma 7.9. *We keep the notation of Lemma 7.8. Then*

$$\left| m(P) - \int_{[0,1)^l} \mu(x)\, dx \right| \ll_{d,k} r^{1/(2k^2)}.$$

Proof. We recall that $|\widehat{P}|\widetilde{P}$ equals $\widehat{P}$ up to a monomial factor. By (7-6), (7-12), and Theorem 2.1 there exists $c > 0$ depending only on k such that $|\widetilde{P}(\boldsymbol{e}(x))| \ge cr^2$ implies $m(P_{\boldsymbol{e}(x)}) \ge \log r$. By Fubini's theorem we have $\int_{[0,1)^l} m(P_{\boldsymbol{e}(x)})\, dx = m(P)$, so the absolute value in question is

$$\mathcal{E} = \left| \int_\Sigma m(P_{\boldsymbol{e}(x)})\, dx - \mathrm{vol}(\Sigma)\log r \right|$$

where $\Sigma = S(\widetilde{P}, cr^2)$; indeed $m(P_{\boldsymbol{e}(x)}) = \mu(x)$ for all $x \in [0,1]^l \setminus \Sigma$.

Note that $\mathrm{vol}(\Sigma) \ll_{d,k} r^{1/(k^2-1)}$ by Lemma A.3(i) applied to $\widetilde{P}$. So

$$\mathcal{E} \ll_{d,k} r^{1/(k^2-1)} |\log r| + \left| \int_\Sigma m(P_{\boldsymbol{e}(x)})\, dx \right| \ll_{d,k} r^{1/k^2} + \int_\Sigma |m(P_{\boldsymbol{e}(x)})|\, dx. \tag{7-13}$$

To bound the integral in (7-13) from above we will replace $m(P_{\boldsymbol{e}(x)})$ by $\log|P_{\boldsymbol{e}(x)}|$. Say $x \in \Sigma$ and $P_{\boldsymbol{e}(x)} \ne 0$, then $\left|m(P_{\boldsymbol{e}(x)}) - \log|P_{\boldsymbol{e}(x)}|\right| \ll_k 1$ by (2-2) and thus $|m(P_{\boldsymbol{e}(x)})| \ll_k 1 + \left|\log|P_{\boldsymbol{e}(x)}|\right|$. The function $x \mapsto \left|\log|P_{\boldsymbol{e}(x)}|\right|$ is integrable over $[0,1)^l$ in the sense of Lebesgue and so is $x \mapsto |m(P_{\boldsymbol{e}(x)})|$; both take the value $+\infty$ on a measure zero subset of $[0,1)^l$. We find

$$\mathcal{E} \ll_{d,k} r^{1/k^2} + \int_\Sigma \left(1 + \left|\log|P_{\boldsymbol{e}(x)}|\right|\right) dx.$$

From (7-6) and (7-12) we deduce $\big|\log|P_{\boldsymbol{e}(x)}|\big| \ll_k \big|\log|\widetilde{P}(\boldsymbol{e}(x))|\big| + 1$ if $\widehat{P}(\boldsymbol{e}(x)) \neq 0$. So $\mathcal{E} \ll_{d,k} r^{1/k^2} + \int_\Sigma \big(1 + \big|\log|\widetilde{P}(\boldsymbol{e}(x))|\big|\big)\,dx$. By Lemma A.4 applied to $\widetilde{P}$ and the volume estimate for Σ, the integral on the right is $\ll_{d,k} r^{1/(k^2-1)} + r^{1/(2(k^2-1))} \ll_{d,k} r^{1/(2k^2)}$, as desired. □

Proof of Proposition 7.3. If we scale P by a factor λ, then $\widehat{P}$, defined in (7-5), is scaled by $|\lambda|^2$. So the proposition is invariant under nonzero scaling and we may assume $|P| = 1$. Later on we will choose the parameter r in terms of $\deg(P)$ and $\mathcal{D}$. In the meantime we assume that $r \in \big(0, \frac{1}{2}\big]$.

We want to bound $\mathcal{E} = |m(P) - n^{-1}\sum_{i=1}^n m(P_{\boldsymbol{e}(x_i)})|$ from above. We replace the Mahler measure with $\mu(\cdot)$ coming from Lemma 7.8. Indeed, the triangle inequality implies

$$\mathcal{E} \le \left|m(P) - \int_{[0,1)^l} \mu(x)\,dx\right| + \left|\int_{[0,1)^l} \mu(x)\,dx - \frac{1}{n}\sum_{i=1}^n \mu(x_i)\right| + \left|\frac{1}{n}\sum_{i=1}^n \mu(x_i) - m(P_{\boldsymbol{e}(x_i)})\right|.$$

The first term on the right is $\ll_{d,k} r^{1/(2k^2)}$ by Lemma 7.9 applied to P. By Proposition 7.1 applied to μ and $t = \mathcal{D}^{1/(d+1)}$ and Lemma 7.8 the second term is $\ll_{d,k} (\deg(P)\mathcal{D}^{1/(d+1)}r^{-1})^{1/(8k)}|\log r|$. So

$$\mathcal{E} \ll_{d,k} r^{1/(2k^2)} + (\deg(P)\mathcal{D}^{1/(d+1)}r^{-2})^{1/(8k)} + \mathcal{E}' \tag{7-14}$$

after absorbing $|\log r|$ in a multiple $r^{-1/(8k)}$ and where $\mathcal{E}'$ is the third term above. Only terms with $m(P_{\boldsymbol{e}(x_i)}) \le \log r$ contribute to the average, so $\mathcal{E}'$ equals

$$\left|\frac{1}{n}\sum_{m(P_{\boldsymbol{e}(x_i)})\le \log r} \log r - m(P_{\boldsymbol{e}(x_i)})\right| \le \frac{|\log r|}{n}\#\{i : m(P_{\boldsymbol{e}(x_i)}) \le \log r\} + \frac{1}{n}\sum_{m(P_{\boldsymbol{e}(x_i)})\le \log r} |m(P_{\boldsymbol{e}(x_i)})|.$$

By Theorem 2.1 we may replace $m(P_{\boldsymbol{e}(x_i)})$ by $\log|P_{\boldsymbol{e}(x_i)}|$ at the cost of introducing a constant $c_1 > 0$ depending only on k, i.e.,

$$\mathcal{E}' \ll_{d,k} \frac{|\log r|}{n}\#\{i : |P_{\boldsymbol{e}(x_i)}| \le c_1 r\} + \frac{1}{n}\sum_{|P_{\boldsymbol{e}(x_i)}|\le c_1 r} \big(1 + \big|\log|P_{\boldsymbol{e}(x_i)}|\big|\big).$$

Recall that $\widetilde{P}$ was defined after the proof of Lemma 7.8. If $|P_{\boldsymbol{e}(x_i)}| \le c_1 r$, then $|\widetilde{P}(\boldsymbol{e}(x_i))| = |\widehat{P}(\boldsymbol{e}(x_i))|/|\widehat{P}| \le c_2 r^2$ for some c_2 depending only on k by (7-6) and (7-12). The same inequalities imply $\big|\log|P_{\boldsymbol{e}(x_i)}|\big| \ll_k \big|\log|\widetilde{P}(\boldsymbol{e}(x_i))|\big| + 1$, the "+1" is absorbed in the first term in

$$\mathcal{E}' \ll_{d,k} \frac{|\log r|}{n}\#\{i : |\widetilde{P}(\boldsymbol{e}(x_i))| \le c_2 r^2\} + \frac{1}{n}\sum_{|\widetilde{P}(\boldsymbol{e}(x_i))|\le c_2 r^2} \big|\log|\widetilde{P}(\boldsymbol{e}(x_i))|\big|.$$

Recall that $\deg \widetilde{P} \ll_d \deg P$ and that $\widetilde{P}$ has at most k^2 terms and norm 1. Lemma 7.4 applied to $\widetilde{P}$ and $c_2 r^2$ implies

$$\mathcal{E}' \ll_{d,k} r^{1/k^2}|\log r| + \deg(P)\mathcal{D}^{1/(l+1)}r^{-2}|\log r| + \frac{1}{n}\sum_{|\widetilde{P}(\boldsymbol{e}(x_i))|\le c_2 r^2} \big|\log|\widetilde{P}(\boldsymbol{e}(x_i))|\big|.$$

We use Lemma 7.7, applied to $\widetilde{P}$ and $c_2 r^2$, to bound the final sum and thus obtain

$$\mathcal{E}' \ll_{d,k} r^{1/k^2} |\log r| + \deg(P)\mathcal{D}^{1/(l+1)} r^{-2} |\log r| + \deg(P)\mathcal{D}^{1/(l+1)} r^{-4} + r^{1/(2k^2)} + \epsilon,$$

here $\epsilon = \left| m(\widehat{P}) - \frac{1}{n}\sum_{i=1}^{n} \log|\widehat{P}(\boldsymbol{e}(x_i))| \right|$; note that multiplying $\widehat{P}$ with a nonzero scalar and a monomial leaves ϵ invariant.

We return to the total error term $\mathcal{E}$. By (7-14) together with $l \le d$, $r \le 1$, and $\mathcal{D} \le 1$ we get

$$\mathcal{E} \ll_{d,k} r^{1/(2k^2)} + (\deg(P)\mathcal{D}^{1/(d+1)} r^{-4})^{1/(8k)} + \deg(P)\mathcal{D}^{1/(d+1)} r^{-4} + \epsilon.$$

We choose $r = \frac{1}{2}\mathcal{D}^{1/(8(d+1))}$, then the proposition follows as $\mathcal{D} \le 1$. □

8. Endgame

In this section we prove a stronger version of Theorem 1.1 from the introduction.

8A. ***Preliminaries.*** Suppose $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}] \backslash \{0\}$. For $V \in \mathrm{GL}_d(\mathbb{Z})$ we set $Q \in \overline{\mathbb{Q}}[X_1, \dots, X_d]$ to be $P(X^{V^{-1}})$ multiplied by a suitable monomial in $X_1, \dots, X_d$ such that Q is coprime to $X_1 \cdots X_d$. Let $l \in \{0, \dots, d-1\}$. For $z = (z_1, \dots, z_l) \in \mathbb{C}^l$ we set

$$P_{V,z} = Q(z_1, \dots, z_l, X_1, \dots, X_{d-l}) \tag{8-1}$$

this is a polynomial in $d - l$ variables. Note that $P_{V,\boldsymbol{e}(x)} = Q_{\boldsymbol{e}(x)} \in \mathbb{C}[X_1, \dots, X_{d-l}]$ in the notation introduced near the beginning of Section 7B. It is useful to allow $l = 0$ in which case $P_{V,z} = Q$. In our typical application $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order. We write $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}, \boldsymbol{\xi})$ where $\boldsymbol{\eta} \in \mathbb{G}_m^l$, $\boldsymbol{\xi} \in \mathbb{G}_m^{d-l}$ and see $|P_{V,\boldsymbol{\eta}}(\boldsymbol{\xi})| = |P(\boldsymbol{\zeta})|$.

The following lemma requires a result of Bombieri, Masser, and Zannier [2007] and relies crucially on the hypothesis that P is essentially atoral.

Lemma 8.1. *Suppose $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm}] \backslash \{0\}$ is essentially atoral. There exists $c \ge 1$ depending only on P and d such that for all $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ of finite order with $\delta(\boldsymbol{\zeta}) \ge c$, for all $V \in \mathrm{GL}_d(\mathbb{Z})$, and all $l \in \{0, \dots, d-1\}$, we have $P_{V,\boldsymbol{\eta}} \ne 0$ and $\mathcal{B}(P_{V,\boldsymbol{\eta}}) \le c|V^{-1}|$ where $\boldsymbol{\zeta}^V = \{\boldsymbol{\eta}\} \times \mathbb{G}_m^{d-l}$.*

Proof. The Zariski closure W in $\mathbb{G}_m^d$ of all algebraic zeros of P in $(S^1)^d$ is defined over $\overline{\mathbb{Q}}$.

By hypothesis, P is essentially atoral. So each irreducible component of the Zariski closure of all complex roots of P on $(S^1)^d$ is of codimension at least 2 in $\mathbb{G}_m^d$ or a proper torsion coset of $\mathbb{G}_m^d$. Therefore, each irreducible component of W is also of this type.

Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ be of finite order with $\delta(\boldsymbol{\zeta}) \ge c$, where c is to be determined, and $\boldsymbol{\zeta}^V = \{\boldsymbol{\eta}\} \times \mathbb{G}_m^{d-l}$ with V and l as in the hypothesis.

Let $\boldsymbol{\eta}' \in \mathbb{G}_m^{d-l}$ be of finite order, $z \in S^1 \backslash \mu_\infty$ be algebraic, and $u \in \mathbb{Z}^{d-l}$ with $P_{V,\boldsymbol{\eta}}(\boldsymbol{\eta}' z^u) = 0$. We must find $v'' \in \mathbb{Z}^{d-l} \backslash \{0\}$ with $|v''| \le c|V^{-1}|$ such that $\langle u, v'' \rangle = 0$. The existence of such a v'' establishes in particular $P_{V,\boldsymbol{\eta}} \ne 0$ (as c depends only on P and d).

Now $P(x) = 0$ for the algebraic point $x = (\boldsymbol{\eta}, \boldsymbol{\eta}' z^u)^{V^{-1}} \in (S^1)^d$. So x is contained in an irreducible component W' of W and in a 1-dimensional algebraic subgroup of $\mathbb{G}_m^d$.

If $\dim W' \leq d-2$, we apply Bombieri, Masser, and Zannier's Theorem 1.5 [2007] to $\mathcal{X} = W'$. We get a proper torsion coset of $\mathbb{G}_m^d$ containing x and coming from a finite set depending only on W', and thus only on P. We find $v \in \mathbb{Z}^d \backslash \{0\}$ with $|v| \ll_{d,P} 1$ and $x^v = 1$.

If W' is a proper torsion coset of $\mathbb{G}_m^d$ there exists $v \in \mathbb{Z}^d \backslash \{0\}$, depending only on W' such that $y^v = 1$ holds for all $y \in W'$. Again we find $|v| \ll_{d,P} 1$ and $x^v = 1$.

In either case we have

$$1 = x^v = (\boldsymbol{\eta}, \boldsymbol{\eta}' z^u)^{V^{-1}v} = \boldsymbol{\eta}^{v'} (\boldsymbol{\eta}' z^u)^{v''} \quad \text{where } V^{-1}v = \begin{pmatrix} v' \\ v'' \end{pmatrix} \in \mathbb{Z}^l \times \mathbb{Z}^{d-l}. \tag{8-2}$$

In particular, $\langle u, v'' \rangle = 0$ as z has infinite order.

If $v'' \neq 0$, then we are done. Indeed, $|v''| \leq |V^{-1}v| \leq d|V^{-1}||v|$ and $|v|$ is bounded from above solely in terms of P and d.

Let us assume $v'' = 0$ and derive a contradiction for c large in terms of P and d. Note $l \geq 1$ as v cannot be 0. Then $v' \neq 0$ and by equality (8-2) we find $\boldsymbol{\eta}^{v'} = 1$. Recall that $\boldsymbol{\eta}$ consists of the first l coordinates of $\boldsymbol{\zeta}^V$. Thus $\boldsymbol{\zeta}^v = 1$ and hence $\delta(\boldsymbol{\zeta}) \leq |v|$ where $|v| \ll_{d,P} 1$. But $\delta(\boldsymbol{\zeta}) \geq c$, a contradiction for large enough c. □

Definition 8.2. Let $c \geq 1$ be a real number. Suppose $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \ldots, X_d^{\pm 1}] \backslash \{0\}$ and $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ is of finite order. The pair $(P, \boldsymbol{\zeta})$ is called c-admissible if for all $V \in \mathrm{GL}_d(\mathbb{Z})$ and all $l \in \{0, \ldots, d-1\}$, we have $P_{V,\boldsymbol{\eta}} \neq 0$ and $\mathcal{B}(P_{V,\boldsymbol{\eta}}) \leq c|V^{-1}|$ where $\boldsymbol{\zeta}^V \in \{\boldsymbol{\eta}\} \times \mathbb{G}_m^{d-l}$.

The case $l = 0$ yields in particular $\mathcal{B}(P) \leq c$ if there exists $\boldsymbol{\zeta}$ such that $(P, \boldsymbol{\zeta})$ is c-admissible; indeed take V as the identity matrix.

Let P be an essentially atoral Laurent polynomial with algebraic coefficient. By Lemma 8.1 there exists $c \geq 1$ such that $(P, \boldsymbol{\zeta})$ is c-admissible for all $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ of finite order with $\delta(\boldsymbol{\zeta}) \geq c$.

In the definition of admissibility, it will be useful to keep track of $\boldsymbol{\zeta}$ when passing it down in an induction step. The next lemma makes this precise.

Lemma 8.3. *Let $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \ldots, X_d^{\pm 1}] \backslash \{0\}$ and let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ be of finite order such that $(P, \boldsymbol{\zeta})$ is c-admissible with $c \geq 1$. Say $l \in \{0, \ldots, d-1\}$, $V \in \mathrm{GL}_d(\mathbb{Z})$, and $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}, \boldsymbol{\xi})$ with $\boldsymbol{\eta} \in \mathbb{G}_m^l$ and $\boldsymbol{\xi} \in \mathbb{G}_m^{d-l}$. Then $(P_{V,\boldsymbol{\eta}}, \boldsymbol{\xi})$ is $(cd|V^{-1}|)$-admissible.*

Proof. Throughout the proof we use that $|\cdot|$ is the maximum-norm on matrices.

We abbreviate $R = P((\boldsymbol{\eta}, X_1, \ldots, X_{d-l})^{V^{-1}})$ which equals $P_{V,\boldsymbol{\eta}}$ up to a monomial factor. It suffices to show that $(R, \boldsymbol{\xi})$ is $(cd|V|^{-1})$-admissible.

To this end say $k \in \{0, \ldots, d-l-1\}$, $W \in \mathrm{GL}_{d-l}(\mathbb{Z})$, and $\boldsymbol{\xi}^W = \{\boldsymbol{\eta}'\} \times \mathbb{G}_m^{d-l-k}$ with $\boldsymbol{\eta}' \in \mathbb{G}_m^k$. We must bound $\mathcal{B}(R_{W,\boldsymbol{\eta}'})$. So say $z \in S^1 \backslash \mu_\infty$, $u \in \mathbb{Z}^{d-l-k}$, and $\boldsymbol{\eta}'' \in \mathbb{G}_m^{d-l-k}$ is of finite order with $R_{W,\boldsymbol{\eta}'}(\boldsymbol{\eta}'' z^u) = 0$. Thus $R((\boldsymbol{\eta}', \boldsymbol{\eta}'' z^u)^{W^{-1}}) = 0$ and hence $P((\boldsymbol{\eta}, (\boldsymbol{\eta}', \boldsymbol{\eta}'' z^u)^{W^{-1}})^{V^{-1}}) = 0$. We abbreviate $W' = \left(\begin{smallmatrix} E_l & 0 \\ 0 & W \end{smallmatrix}\right)$ with E_l the $l \times l$ identity matrix. So $P((\boldsymbol{\eta}, \boldsymbol{\eta}', \boldsymbol{\eta}'' z^u)^{(VW')^{-1}}) = 0$ which means $P_{VW',(\boldsymbol{\eta},\boldsymbol{\eta}')}(\boldsymbol{\eta}'' z^u) = 0$.

Observe that $\boldsymbol{\zeta}^{VW'} = (\boldsymbol{\eta}, \boldsymbol{\xi})^{W'} = (\boldsymbol{\eta}, \boldsymbol{\xi}^W) = (\boldsymbol{\eta}, \boldsymbol{\eta}', *)$. By hypothesis $(P, \boldsymbol{\zeta})$ is c-admissible. Therefore, $\mathcal{B}(P_{VW',(\boldsymbol{\eta},\boldsymbol{\eta}')}) \leq c|(VW')^{-1}| = c|W'^{-1}V^{-1}| \leq cd|V^{-1}||W'^{-1}| = cd|V^{-1}||W^{-1}|$. In other words, there

exists $v \in \mathbb{Z}^{d-l-k}\backslash\{0\}$ with $|v| \le cd|V^{-1}||W^{-1}|$ and $\langle u, v\rangle = 0$. Thus $\mathcal{B}(R_{W,\boldsymbol{\eta}'}) \le cd|V^{-1}||W^{-1}|$, as desired. Moreover, $R_{W,\boldsymbol{\eta}'} \neq 0$. □

Lemma 8.4. *Let $P \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]\backslash\{0\}$ and let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ be of finite order such that $(P, \boldsymbol{\zeta})$ is c-admissible with $c \ge 1$. Say $l \in \{1, \dots, d-1\}$ and let $\widehat{P} \in \overline{\mathbb{Q}}[X_1^{\pm 1}, \dots, X_l^{\pm 1}]$ be as in (7-5) and $\boldsymbol{\zeta} \in \{\boldsymbol{\eta}\} \times \mathbb{G}_m^{d-l}$. Then $(\widehat{P}, \boldsymbol{\eta})$ is c-admissible.*

Proof. Suppose $V \in \mathrm{GL}_l(\mathbb{Z})$ such that $\boldsymbol{\eta}^V = (\boldsymbol{\eta}', *)$ where $\boldsymbol{\eta} \in \mathbb{G}_m^{l'}$ and $l' \in \{0, \dots, l-1\}$. Following the definition of admissibility and recalling (8-1) we are in the following situation. There is $\boldsymbol{\eta}'' \in \mathbb{G}_m^{l-l'}$, $z \in S^1\backslash\mu_\infty$ algebraic, and $u' \in \mathbb{Z}^{l-l'}$ such that

$$\widehat{P}((\boldsymbol{\eta}', \boldsymbol{\eta}'' z^{u'})^{V^{-1}}) = 0.$$

It follows from the definition of $\widehat{P}$ that $P((\boldsymbol{\eta}', \boldsymbol{\eta}'' z^{u'})^{V^{-1}}, X_{l+1}, \dots, X_d) = 0$ as a polynomial in $X_{l+1}, \dots, X_d$. We extend $\widetilde{V} = \left(\begin{smallmatrix} V & 0 \\ 0 & E_{d-l} \end{smallmatrix}\right)$ where E_{d-l} is the $(d-l) \times (d-l)$ identity matrix. Then $P((\boldsymbol{\eta}', \boldsymbol{\eta}'' z^{u'}, z^{u''})^{\widetilde{V}^{-1}}) = 0$ for all $u'' \in \mathbb{Z}^{d-l}$.

By hypothesis, $(P, \boldsymbol{\zeta})$ is c-admissible and $\boldsymbol{\zeta}^{\widetilde{V}} = (\boldsymbol{\eta}^V, *) = (\boldsymbol{\eta}', *, *)$. Now $P_{\widetilde{V},\boldsymbol{\eta}'}(\boldsymbol{\eta}'' z^{u'}, z^{u''}) = 0$, so by definition there exist $v' \in \mathbb{Z}^{l-l'}$, $v'' \in \mathbb{Z}^{d-l}$, not both zero, such that $\langle u', v'\rangle + \langle u'', v''\rangle = 0$ and $|(v', v'')| \le c|\widetilde{V}^{-1}| = c|V^{-1}|$ for the maximum-norm.

As we are free to vary u'' we see that $\{u'\} \times \mathbb{Q}^{d-l}$ is contained in a finite union of proper vector subspaces of $\mathbb{Q}^d$, each defined as the kernel of $\langle \cdot, (v', v'')\rangle$ with v', v'' as above. So $\{u'\} \times \mathbb{Q}^{d-l} \subset V$ for one of these vector spaces V defined by some (v', v''). We must have $v'' = 0$ and hence $\langle u', v'\rangle = 0$. Then $v' \neq 0$ and as $|v'| \le c|V^{-1}|$ we conclude that $\widehat{P}$ is c-admissible. □

Here are some basic estimates involving $P_{V,\boldsymbol{\eta}}$.

Lemma 8.5. *Let $P \in \overline{\mathbb{Q}}[X_1, \dots, X_d]\backslash\{0\}$, $l \in \{0, \dots, d-1\}$, and $V \in \mathrm{GL}_d(\mathbb{Z})$. Say $\boldsymbol{\eta} \in \mathbb{G}_m^l$ has finite order and $P_{V,\boldsymbol{\eta}} \neq 0$: The following hold true.*

(i) *We have* $\deg P_{V,\boldsymbol{\eta}} \ll_d |V|^{d-1} \deg P$.

(ii) *We have $h(P_{V,\boldsymbol{\eta}}) \le \log(k) + h(P)$ where $k \ge 2$ is an upper bound for the number of nonzero terms of P.*

Proof. Both parts follow are elementary consequences of the degree and the height of a polynomial. For (i) we require $|V^{-1}| \ll_d |V|^{d-1}$. For (ii) we note that Q from the beginning of this subsection has the same coefficients and thus the same height as P. We decompose $h(P_{V,\boldsymbol{\eta}})$ in local heights as in (2-4). The triangle inequality at the archimedean places leads to $\log k$. □

We continue with further basic estimates involving $\widehat{P}$ as in (7-5).

Lemma 8.6. *Let $K \subset \mathbb{C}$ be a number field and suppose $P \in K[X_1, \dots, X_d]\backslash\{0\}$ has at most $k \ge 2$ terms, where k is an integer. Say $l \in \{1, \dots, d-1\}$ with $\widehat{P} \in \mathbb{C}[X_1^{\pm 1}, \dots, X_l^{\pm 1}]$ as in (7-5). Then the following properties hold true*:

(i) *We have* $\widehat{P} \in K'[X_1^{\pm 1}, \dots, X_l^{\pm 1}]$ *where* K' *is a number field such that* $K \subset K' \subset \mathbb{C}$ *and* $[K' : \mathbb{Q}] \le [K : \mathbb{Q}]^2$.

(ii) *We have* $h(\widehat{P}) \ll_k 1 + h(P)$.

Proof. The coefficients of $\widehat{P}$ are contained in the subfield K' of $\mathbb{C}$ generated by a primitive element of $K/\mathbb{Q}$ and its complex conjugate. So $[K' : \mathbb{Q}] \le [K : \mathbb{Q}]^2$ and (i) follows. For (ii) we remark that each p_i as in (7-5) has at most k terms and that there are at most k nonzero p_i. Using the local decomposition of the height together with the ultrametric and archimedean triangle inequality yields the claim. □

8B. ***Completion of proof.*** The next lemma will setup a monomial change of coordinates. We recall that $\Lambda_{\boldsymbol{\xi}}$ was defined in (5-4), $\widetilde{\Lambda}_{\boldsymbol{\xi}}(\nu)$ was defined in (5-5) and $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}}(\nu))$ is as in (5-6).

Lemma 8.7. *Suppose* $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ *has order* N *and let* $\delta \ge 1$, $\epsilon \in \left(0, \frac{1}{2}\right]$, $\nu_1, \dots, \nu_{d-1} \in \left(0, \frac{1}{2}\right]$ *with* $\nu_1 + \cdots + \nu_{d-1} \le \frac{1}{2}$. *Then there exist* $l \in \{0, \dots, d-1\}$ *and* $V \in \mathrm{GL}_d(\mathbb{Z})$ *such that the following hold*:

(i) *We have* $|V| \ll_d \delta^{2\epsilon^{d-l}}$ *and* V *is the identity matrix if* $l = 0$.

(ii) *We have* $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}, \boldsymbol{\xi})$ *where* $\boldsymbol{\eta} \in \mathbb{G}_m^l$, $\boldsymbol{\xi} \in \mathbb{G}_m^{d-l}$, $\mathrm{ord}(\boldsymbol{\eta}) \le N^{\nu_1 + \cdots + \nu_l}$, $\boldsymbol{\xi}$ *has finite order at least* $N^{1/2}$. *Finally, if* $l \le d-2$ *then* $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}}(\nu_{l+1})) > \delta^{\epsilon^{d-l-1}}$.

Proof. Set $\boldsymbol{\xi}_1 = \boldsymbol{\zeta}$ and let V_0 be the identity matrix in $\mathrm{GL}_d(\mathbb{Z})$. For all $l \in \{1, \dots, d-1\}$ with $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l)) \le \delta^{\epsilon^{d-l}}$ we will construct inductively $V_l \in \mathrm{GL}_d(\mathbb{Z})$, $\boldsymbol{\xi}_{l+1} \in \mathbb{G}_m^{d-l}$ of order at most N, and $\eta_l \in \mathbb{G}_m$ of order at most N^{ν_l} such that $\boldsymbol{\zeta}^{V_l} = (\eta_1, \dots, \eta_l, \boldsymbol{\xi}_{l+1})$ and

$$|V_l| \ll_d \delta^{\epsilon^{d-1} + \cdots + \epsilon^{d-l}}. \tag{8-3}$$

Suppose $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l)) \le \delta^{\epsilon^{d-l}}$, there exists $v \in \widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l) \backslash \{0\}$ such that $|v| \le \delta^{\epsilon^{d-l}}$ and v is primitive. Note that $[\widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l) : \Lambda_{\boldsymbol{\xi}_l}(\nu_l)]v$ lies in $\Lambda_{\boldsymbol{\xi}_l}$, so $\mathrm{ord}(\boldsymbol{\xi}_l^v) \le [\widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l) : \Lambda_{\boldsymbol{\xi}_l}(\nu_l)] \le \det(\Lambda_{\boldsymbol{\xi}_l}(\nu_l)) \le N^{2\nu_l^2} \le N^{\nu_l}$ by (5-3) and since $\det(\Lambda_{\boldsymbol{\xi}_l}) = \mathrm{ord}(\boldsymbol{\xi}_l) \le N$. We can realize v as the first column of a matrix $V_l' \in \mathrm{GL}_{d-l+1}(\mathbb{Z})$ with $|V_l'| \ll_d |v| \ll_d \delta^{\epsilon^{d-l}}$, see the proof of Proposition 5.1. Let E_{l-1} denote the $(l-1) \times (l-1)$ identity matrix and set

$$V_l = V_{l-1} \begin{pmatrix} E_{l-1} & 0 \\ 0 & V_l' \end{pmatrix} \in \mathrm{GL}_d(\mathbb{Z}).$$

By step $l-1$ we have $\boldsymbol{\zeta}^{V_{l-1}} = (\eta_1, \dots, \eta_{l-1}, \boldsymbol{\xi}_l)$. We define η_l and $\boldsymbol{\xi}_{l+1}$ via $\boldsymbol{\zeta}^{V_l} = (\eta_1, \dots, \eta_l, \boldsymbol{\xi}_{l+1})$. Note $\eta_l = \boldsymbol{\xi}_l^v$, so $\mathrm{ord}(\eta_l) \le N^{\nu_l}$ by the bound above. Finally, $|V_l| \ll_d |V_{l-1}||V_l'| \ll_d \delta^{\epsilon^{d-1} + \cdots + \epsilon^{d-l}}$, and $\boldsymbol{\xi}_l^N = 1$. This completes our construction.

Otherwise, fix the largest $l \in \{1, \dots, d-1\}$ for which $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}_l}(\nu_l)) \le \delta^{\epsilon^{d-l}}$; if no l satisfies the inequality we take $l = 0$. Then define $V = V_l$. Thus V is the identity matrix if $l = 0$ and claims (i) and (ii) are immediate as $\boldsymbol{\xi} = \boldsymbol{\zeta}$. So say $l \ge 1$. Then (i) holds by (8-3) as $\epsilon \le \frac{1}{2}$. To verify (ii) observe that $\boldsymbol{\zeta}^V = (\eta_1, \dots, \eta_l, \boldsymbol{\xi})$ with $\boldsymbol{\xi} = \boldsymbol{\xi}_{l+1} \in \mathbb{G}_m^{d-l}$ and $(\eta_1, \dots, \eta_l)$ has order at most $N^{\nu_1 + \cdots + \nu_l} \le N^{1/2}$. Thus $\boldsymbol{\xi}$ has finite order at least $N^{1/2}$ since $\boldsymbol{\zeta}^V$ has order N. If $l \le d-2$, then $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}}(\nu_{l+1})) > \delta^{\epsilon^{d-l-1}}$, because the construction does not continue. □

We are ready to prove a theorem that will quickly imply our Theorem 1.1 and its refinements.

Theorem 8.8. *Let $c \geq 1$, let $K \subset \mathbb{C}$ be a number field, and suppose $P \in K[X_1, \dots, X_d]\backslash\{0\}$ has at most k terms for an integer $k \geq 2$. There are constants $C = C(d,k) \geq 1$ and $\kappa = \kappa(d,k) > 0$ depending only on d and k with the following property. Let $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ have finite order N and suppose G is a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ with $P(\boldsymbol{\zeta}^\sigma) \neq 0$ for all $\sigma \in G$. If $(P, \boldsymbol{\zeta}^\sigma)$ is c-admissible for all $\sigma \in G$ and if*

$$\delta(\boldsymbol{\zeta}) \geq C \max\{c, \deg P\}^C \tag{8-4}$$

then

$$\frac{1}{\#G}\sum_{\sigma\in G} \log|P(\boldsymbol{\zeta}^\sigma)| = m(P) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^{2^d}[(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^{16d^2}(1+h(P))}{\delta(\boldsymbol{\zeta})^\kappa}\right).$$

Proof. The case $d = 1$ follows from Proposition 6.2(i) as $\delta(\boldsymbol{\zeta}) = N$ in this case and as $\mathcal{B}(P) < \infty$. So we may assume $d \geq 2$. We may also assume that P is nonconstant.

We work with the parameters $\nu_1, \dots, \nu_{d-1} \in (0, 1/(128d^2)]$, $\epsilon \in \left(0, \frac{1}{2}\right]$ in this proof. They are assumed to be small in terms of d and k but independent of P and $\boldsymbol{\zeta}$. We may assume that ϵ is small in terms of the ν_l, e.g., $\epsilon \leq \nu_l^d/4$ for all l. We determine them during the argument.

We apply Lemma 8.7 to $\boldsymbol{\zeta}$, $\delta = \delta(\boldsymbol{\zeta})$, ϵ, and the ν_l. Say l, V, $\boldsymbol{\eta}$, and $\boldsymbol{\xi}$ are given by this lemma, in particular $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}, \boldsymbol{\xi})$ and $|V| \ll_d \delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}}$. We have

$$\operatorname{ord}(\boldsymbol{\eta}) \leq N^{\nu_1+\cdots+\nu_l} \quad \text{and} \quad \operatorname{ord}(\boldsymbol{\xi}) \geq N^{1/2}. \tag{8-5}$$

The case $l = 0$ is straightforward. Here V is the identity matrix, $\boldsymbol{\xi} = \boldsymbol{\zeta}$, and $\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\zeta}}(\nu_1)) > \delta(\boldsymbol{\zeta})^{\epsilon^{d-1}}$ as we are in case $d - l = d \geq 2$ of Lemma 8.7(ii). So $\tilde{\lambda}(\boldsymbol{\zeta}; \nu_1) \geq \delta(\boldsymbol{\zeta})^{\min\{\epsilon^{d-1}, \nu_1^d/2\}}$ using (6-2) and $\delta(\boldsymbol{\zeta}) \leq N$. As $(P, \boldsymbol{\zeta})$ is c-admissible we have $\mathcal{B}(P) \leq c$. We will apply Proposition 6.2(ii) to P and $\nu = \nu_1$, so we must verify $\tilde{\lambda}(\boldsymbol{\zeta}; \nu_1) > d^{1/2}\max\{c, \deg P\}$. This inequality is satisfied if $\delta(\boldsymbol{\zeta})$ is as in (8-4) with C large in terms of ϵ, ν_1, d, and k. So if $l = 0$, the theorem follows from (6-3).

Step 1: A monomial change of coordinates. From now on we assume $l \geq 1$, i.e., $l \in \{1, \dots, d-1\}$. We fix $\sigma \in G$ throughout this set. We have $\boldsymbol{\zeta}^{\sigma V} = (\boldsymbol{\eta}^\sigma, \boldsymbol{\xi}^\sigma) \in \mathbb{G}_m^l \times \mathbb{G}_m^{d-l}$ and $|P(\boldsymbol{\zeta}^\sigma)| = |P_{V,\boldsymbol{\eta}^\sigma}(\boldsymbol{\xi}^\sigma)|$. This time we apply Proposition 6.2 to $P_{V,\boldsymbol{\eta}^\sigma} \in K(\boldsymbol{\eta})[X_1, \dots, X_{d-l}]$, $\boldsymbol{\xi}^\sigma$, and ν_{l+1}. We often use that $\delta(\cdot)$ is Galois invariant, i.e., $\delta(\boldsymbol{\zeta}^\sigma) = \delta(\boldsymbol{\zeta})$.

If $d - l = 1$ we will apply Proposition 6.2(i) and there is nothing further to check.

But for $d - l \geq 2$ we must verify the hypothesis in the second part of this proposition. This step is similar as in the case $l = 0$.

Note that $P_{V,\boldsymbol{\eta}^\sigma} \neq 0$ as $(P, \boldsymbol{\zeta}^\sigma)$ is c-admissible; this polynomial has at most k nonzero terms. By Lemma 8.3 the pair $(P_{V,\boldsymbol{\eta}^\sigma}, \boldsymbol{\xi}^\sigma)$ is $(cd|V^{-1}|)$-admissible. Observe $|V^{-1}| \ll_d |V|^{d-1} \ll_d \delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}(d-1)}$. So the said pair is $c_1 c\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}$-admissible; here and below $c_1, c_2, \dots$ denote positive constants that depend only on d. In particular, $\mathcal{B}(P_{V,\boldsymbol{\eta}^\sigma}) \leq c_1 c\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}$. A similar argument and Lemma 8.5(i) yield

$$\max\{\mathcal{B}(P_{V,\boldsymbol{\eta}^\sigma}), \deg P_{V,\boldsymbol{\eta}^\sigma}\} \leq c_2\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}\max\{c, \deg P\}.$$

In conclusion, to apply Proposition 6.2(ii) we must verify the inequality in

$$\tilde{\lambda}(\boldsymbol{\xi}^\sigma; \nu_{l+1}) = \max\{\lambda_1(\widetilde{\Lambda}_{\boldsymbol{\xi}^\sigma}(\nu_{l+1})), \operatorname{ord}(\boldsymbol{\xi}^\sigma)^{\nu_{l+1}^{d-l}/2}\} > c_2\sqrt{d}\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}\max\{c, \deg P\}.$$

But $\Lambda_{\boldsymbol{\xi}^\sigma} = \Lambda_{\boldsymbol{\xi}}$ and the order are Galois invariant, see (5-4) and (5-5). So it suffices to prove the lower bound for $\tilde{\lambda}(\boldsymbol{\xi}; \nu_{l+1})$. By Lemma 8.7(ii) we have

$$\tilde{\lambda}(\boldsymbol{\xi}; \nu_{l+1}) \geq \min\{\delta(\boldsymbol{\zeta})^{\epsilon^{d-l-1}}, N^{\nu_{l+1}^{d-l}/4}\} \geq \delta(\boldsymbol{\zeta})^{\min\{\epsilon^{d-l-1}, \nu_{l+1}^d/4\}} = \delta(\boldsymbol{\zeta})^{\epsilon^{d-l-1}} \tag{8-6}$$

as $\epsilon \leq \nu_{l+1}^d/4$. We may assume $\epsilon^{d-l-1} - 2\epsilon^{d-l}d \geq \epsilon^{d-l-1}/2$; this is equivalent to $\epsilon \leq 1/(4d)$. By (8-4) and (8-6) the desired inequality is satisfied when C is large in terms of ϵ, d, and k. We may thus apply Proposition 6.2.

We collect the following bounds from Lemmas 8.5 and 8.7:

$$\begin{aligned} \deg P_{V,\boldsymbol{\eta}^\sigma} &\ll_{d,k} \delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}\deg P, \\ h(P_{V,\boldsymbol{\eta}^\sigma}) &\ll_k 1 + h(P), \\ [K(\boldsymbol{\eta}) : \mathbb{Q}] \leq \operatorname{ord}(\boldsymbol{\eta})[K : \mathbb{Q}] &\leq N^{\nu_1+\cdots+\nu_l}[K : \mathbb{Q}], \text{ and} \\ \operatorname{ord}(\boldsymbol{\xi}) &\geq N^{1/2} \end{aligned} \tag{8-7}$$

Recall that $\boldsymbol{\zeta}^V = (\boldsymbol{\eta}, \boldsymbol{\xi})$. So $\boldsymbol{\xi}^u = 1$ for some $u \in \mathbb{Z}^{d-l}$ with $|u| = \delta(\boldsymbol{\xi})$, hence $\boldsymbol{\zeta}^{V\binom{0}{u}} = 1$. We conclude $\delta(\boldsymbol{\zeta}) \leq \left|V\binom{0}{u}\right| \leq d|V|\delta(\boldsymbol{\xi})$. We find

$$\delta(\boldsymbol{\xi}^\sigma) = \delta(\boldsymbol{\xi}) \gg_d \delta(\boldsymbol{\zeta})^{1-2\epsilon^{d-l}} \gg_d \delta(\boldsymbol{\zeta})^{1/2} \tag{8-8}$$

as we may assume $\epsilon^{d-l} \leq \frac{1}{4}$.

We must specify a subgroup $G \subset (\mathbb{Z}/M\mathbb{Z})^\times$ in Proposition 6.2 where $M = \operatorname{ord}(\boldsymbol{\xi})$; we will denote it by H here. Let L denote the fixed field of G in $\mathbb{Q}(\boldsymbol{\zeta})$. Let H be the subgroup of $(\mathbb{Z}/M\mathbb{Z})^\times = \operatorname{Gal}(\mathbb{Q}(\boldsymbol{\xi})/\mathbb{Q})$ corresponding to $\operatorname{Gal}(\mathbb{Q}(\boldsymbol{\xi})/\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}))$. We identify H with $\operatorname{Gal}(L(\boldsymbol{\zeta})/L(\boldsymbol{\eta}))$ under the isomorphism $\operatorname{Gal}(L(\boldsymbol{\zeta})/L(\boldsymbol{\eta})) \to \operatorname{Gal}(\mathbb{Q}(\boldsymbol{\xi})/\mathbb{Q}(\boldsymbol{\xi}) \cap L(\boldsymbol{\eta}))$ induced by restriction. The fixed field of H in $\mathbb{Q}(\boldsymbol{\xi})$ is contained in $L(\boldsymbol{\eta})$, so

$$[(\mathbb{Z}/M\mathbb{Z})^\times : H] \leq [L(\boldsymbol{\eta}) : \mathbb{Q}] \leq \operatorname{ord}(\boldsymbol{\eta})[L : \mathbb{Q}] = \operatorname{ord}(\boldsymbol{\eta})[(\mathbb{Z}/N\mathbb{Z})^\times : G] \leq N^{\nu_1+\cdots+\nu_l}[(\mathbb{Z}/N\mathbb{Z})^\times : G]$$

having used the bound for the order of $\boldsymbol{\eta}$ from (8-5). Moreover, the conductor of H satisfies

$$f_H \leq \operatorname{lcm}(f_G, \operatorname{ord}(\boldsymbol{\eta})) \leq f_G \operatorname{ord}(\boldsymbol{\eta}) \leq f_G N^{\nu_1+\cdots+\nu_l}.$$

If $\tau \in H$, then $\boldsymbol{\eta}^\tau = \boldsymbol{\eta}$. Therefore, $|P(\boldsymbol{\zeta}^{\tau\sigma})| = |P_{V,\boldsymbol{\eta}^{\tau\sigma}}(\boldsymbol{\xi}^{\tau\sigma})| = |P_{V,\boldsymbol{\eta}^\sigma}(\boldsymbol{\xi}^{\tau\sigma})| \neq 0$ for all $\tau \in H$. To cover the case $l = d-1$ it is useful to set $\nu_d = 1/(128d^2)$. By applying Proposition 6.2 to $P_{V,\boldsymbol{\eta}^\sigma}, \boldsymbol{\xi}^\sigma, \nu_{l+1}$,

and H while using the various estimates above, include (8-6) and (8-7), we find

$$\begin{aligned}\frac{1}{\#H}\sum_{\tau\in H}\log|P_{V,\boldsymbol{\eta}^\sigma}(\boldsymbol{\xi}^{\tau\sigma})|&\\ =\frac{1}{\#H}\sum_{\tau\in H}\log|P(\boldsymbol{\zeta}^{\tau\sigma})|&\\ =m(P_{V,\boldsymbol{\eta}^\sigma})+O_{d,k}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times:G]^2 f_G\deg(P)^2(1+h(P))N^{(2+2+1)(\nu_1+\cdots+\nu_l)}\delta(\boldsymbol{\zeta})^{4\epsilon^{d-l}d}}{N^{\nu_{l+1}^d/(40d)}}\right)&\\ +O_{d,k}\left(\frac{\deg(P)^{16d^2}}{\delta(\boldsymbol{\zeta})^{\epsilon^{d-l-1}/(16k)-32\epsilon^{d-l}d^3}}\right)&\end{aligned}$$

here we used $r\le d$ and $M\ge N^{1/2}$; the second error term, which appears in the their line of the expression, can be omitted if $l=d-1$ as then we apply Proposition 6.2(i).

At this point we reap the benefit of having split the error term in Proposition 6.2 into a part depending on N and a part depending on $\delta(\boldsymbol{\zeta})$. Indeed, the order of $\boldsymbol{\eta}$, which we bound in terms of N, does not affect the second error term above. Recall that $\delta(\boldsymbol{\zeta})\le N$, but there can be no meaningful lower bound for $\delta(\boldsymbol{\zeta})$ in terms of N. Introducing a dependency on N in the second error term $\delta(\boldsymbol{\zeta})$ would spoil the result.

We use the crude bound $\delta(\boldsymbol{\zeta})\le N$ and we may assume the parameters satisfy

$$5(\nu_1+\cdots+\nu_l)+4\epsilon^{d-l}d\le\frac{\nu_{l+1}^d}{80d},$$

for all $l\in\{1,\dots,d-1\}$, and

$$32\epsilon^{d-l}d^3\le\frac{\epsilon^{d-l-1}}{32k}.$$

Such a choice is possible. Indeed, we may fix $\nu_d,\nu_{d-1},\dots,\nu_1$ to decay quickly enough and ϵ is allowed to be small in terms of $\nu_1,\dots,\nu_{d-1}$ and d,k.

We now combine both contributions to the error term and get

$$\frac{1}{\#H}\sum_{\tau\in H}\log|P(\boldsymbol{\zeta}^{\tau\sigma})|=m(P_{V,\boldsymbol{\eta}^\sigma})+O_{d,k}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^\times:G]^2 f_G\deg(P)^{16d^2}(1+h(P))}{\delta(\boldsymbol{\zeta})^\kappa}\right)\tag{8-9}$$

if

$$\kappa\le\min\left\{\frac{\nu_{l+1}^d}{80},\frac{\epsilon^{d-l-1}}{32k}\right\}.$$

Later we may shrink κ.

Step 2: Induction on d. Let $\boldsymbol{\zeta}$ be as in the hypothesis. Recall that $\boldsymbol{\zeta}^V=(\boldsymbol{\eta},\boldsymbol{\xi})$. We still assume $l\ge1$ and we find that

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^\sigma)|=\frac{1}{[L(\boldsymbol{\eta}):L]}\sum_{\tau\in\mathrm{Gal}(L(\boldsymbol{\eta})/L)}\frac{1}{\#H}\sum_{\sigma\in H}\log|P(\boldsymbol{\zeta}^{\tilde\tau\sigma})|$$

with $\tilde{\tau}$ a fixed lift of τ to $\mathrm{Gal}(L(\boldsymbol{\zeta})/L) = \mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/L) = G$, recall that $H = \mathrm{Gal}(L(\boldsymbol{\zeta})/L(\boldsymbol{\eta}))$. Thus (8-9) with $\tilde{\tau}$ for σ implies

$$\frac{1}{\#G}\sum_{\sigma\in G}\log|P(\boldsymbol{\zeta}^{\sigma})| = \frac{1}{[L(\boldsymbol{\eta}):L]}\sum_{\tau\in\mathrm{Gal}(L(\boldsymbol{\eta})/L)} m(P_{V,\boldsymbol{\eta}^{\tau}}) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^2[(\mathbb{Z}/N\mathbb{Z})^{\times}:G]^2 f_G \deg(P)^{16d^2}(1+h(P))}{\delta(\boldsymbol{\zeta})^{\kappa}}\right). \tag{8-10}$$

We set Q to equal $P(X^{V^{-1}})$ times a monomial such that Q is a polynomial coprime to $X_1\cdots X_d$. We apply the construction (7-5) to Q and l and obtain $\widehat{Q}$. Recall Lemma 7.2 and write $\widetilde{Q}$ for $\widehat{Q}$ times the monomial from part (ii) of this lemma. Then $\widetilde{Q}$ has at most k^2 nonzero terms and using also Lemma 8.6 we find

$$\begin{aligned} &\widetilde{Q}\in K'[X_1^{\pm1},\dots,X_l^{\pm1}] \quad \text{where } [K':\mathbb{Q}]\le[K:\mathbb{Q}]^2,\\ &\deg\widetilde{Q}\ll_d \deg Q\ll_d |V^{-1}|\deg P\ll_d |V|^{d-1}\deg P\ll_d \delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}\deg P, \text{ and}\\ &h(\widetilde{Q})\ll_k 1+h(Q)\ll_k 1+h(P). \end{aligned} \tag{8-11}$$

By Lemma 8.3, with $l=0$, the pair $(Q,(\boldsymbol{\eta}^{\sigma},\boldsymbol{\xi}^{\sigma}))$ is $c_3c\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}$-admissible for all $\sigma\in G$. Now $(\widetilde{Q},\boldsymbol{\eta}^{\sigma})$ is also $c_3c\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d}$-admissible by Lemma 8.4 for all $\sigma\in G$ (multiplying a polynomial by a monomial has no effect on admissibility).

We want to apply the current theorem to $\widetilde{Q}$ and $\boldsymbol{\eta}\in\mathbb{G}_m^l$ by induction on the number of variables, recall $l\le d-1$. For this we must verify

$$\delta(\boldsymbol{\eta})\ge c_4C(l,k^2)\delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}dC(l,k^2)}\max\{c,\deg P\}^{C(l,k^2)}$$

having used the bound for $\deg\widetilde{Q}$ in (8-11). As above and in (8-8), the bound $\delta(\boldsymbol{\eta})\ge\delta(\boldsymbol{\zeta})/|V|\gg_d\delta(\boldsymbol{\zeta})^{1/2}$. So it suffices to check

$$\delta(\boldsymbol{\zeta})^{1-4\epsilon^{d-l}dC(l,k^2)}\ge c_5C(l,k^2)^2\max\{c,\deg P\}^{2C(l,k^2)}. \tag{8-12}$$

We may assume that $1-4\epsilon^{d-l}dC(l,k^2)\ge\frac{1}{2}$ as we may choose ϵ small in terms of d and $C(l,k^2)$. So (8-12) follows from (8-4) if $C=C(d,k)$ is large enough in terms of d and k.

To apply this theorem by induction we must specify a subgroup of $H'\subset(\mathbb{Z}/E\mathbb{Z})^{\times}=\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\eta})/\mathbb{Q})$ with $E=\mathrm{ord}(\boldsymbol{\eta})$. We take H' as identified with $\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\eta})/\mathbb{Q}(\boldsymbol{\eta})\cap L)\cong\mathrm{Gal}(L(\boldsymbol{\eta})/L)$ under the isomorphism induced by restriction. For all $\tau\in\mathrm{Gal}(L(\boldsymbol{\eta})/L)$ we have $P_{V,\boldsymbol{\eta}^{\tau}}\ne0$ and so $\widetilde{Q}(\boldsymbol{\eta}^{\tau})\ne0$ by (7-6). By induction and (8-11) we have

$$\frac{1}{\#H'}\sum_{\tau\in H'}\log|\widetilde{Q}(\boldsymbol{\eta}^{\tau})| = m(\widetilde{Q}) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^{2^d}[(\mathbb{Z}/E\mathbb{Z})^{\times}:H']^2 f_{H'}\deg(P)^{16d^2}\delta(\boldsymbol{\zeta})^{32\epsilon^{d-l}d^3}(1+h(P))}{\delta(\boldsymbol{\eta})^{\kappa(l,k^2)}}\right).$$

Note that $[(\mathbb{Z}/E\mathbb{Z})^\times : H'] = [\mathbb{Q}(\boldsymbol{\eta}) \cap L : \mathbb{Q}] \le [L : \mathbb{Q}] = [(\mathbb{Z}/N\mathbb{Z})^\times : G]$ and $f_{H'} \le f_G$. Using again $\delta(\boldsymbol{\eta}) \gg_d \delta(\boldsymbol{\zeta})^{1/2}$ we get

$$\frac{1}{\#H'} \sum_{\tau \in H'} \log|\widehat{Q}(\boldsymbol{\eta}^\tau)| = m(\widehat{Q}) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^{2^d} [(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^{16d^2} (1+h(P))}{\delta(\boldsymbol{\zeta})^{\kappa(l,k^2)/2 - 32\epsilon^{d-l}d^3}} \right) \tag{8-13}$$

as passing from $\widetilde{Q}$ to $\widehat{Q}$ is harmless. We may assume that $\kappa(l,k^2)/4 \ge 32\epsilon^{d-l}d^3$.

Recall that Q equals $P(X^{V^{-1}})$ up to a monomial factor. We will soon apply Proposition 7.3 to Q. Consider $(x_1, \dots, x_{\#H'})$, with each $x_i \in [0,1)^l$, a tuple of discrepancy $\mathcal{D}$ as in (3-2), where the $\boldsymbol{e}(x_i)$ are the $\boldsymbol{\eta}^\tau$. So $P_{V,\boldsymbol{\eta}^\tau} = P_{V,\boldsymbol{e}(x_i)} = Q_{\boldsymbol{e}(x_i)}$. Proposition 7.3 together with (8-13) imply

$$\begin{aligned}\frac{1}{\#H'} \sum_{\tau \in H'} & m(P_{V,\boldsymbol{\eta}^\tau}) \\ &= m(Q) + O_{d,k}\left(\deg(Q)\mathcal{D}^{1/(16(d+1)k^2)} + \frac{[K:\mathbb{Q}]^{2^d} [(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^{16d^2} (1+h(P))}{\delta(\boldsymbol{\zeta})^{\kappa(l,k^2)/4}} \right).\end{aligned}$$

By Proposition 3.3(i) for $\boldsymbol{\eta}$ and H' and estimates used above we find

$$\mathcal{D} \ll_d [(\mathbb{Z}/E\mathbb{Z})^\times : H'] f_{H'}^{1/2} \delta(\boldsymbol{\eta})^{-1/3} \ll_d [(\mathbb{Z}/N\mathbb{Z})^\times : G] f_G^{1/2} \delta(\boldsymbol{\zeta})^{-1/6}.$$

From above we find $\deg Q \ll_d |V^{-1}| \deg P \ll_d \delta(\boldsymbol{\zeta})^{2\epsilon^{d-l}d} \deg P$. The Mahler measure is invariant under a monomial change of coordinates by [Schinzel 2000, Corollary 8, Chapter 3.4], thus $m(P) = m(Q)$. As H' is identified with $\mathrm{Gal}(L(\boldsymbol{\eta})/L)$ we get

$$\frac{1}{[L(\boldsymbol{\eta}) : L]} \sum_{\tau \in \mathrm{Gal}(L(\boldsymbol{\eta})/L)} m(P_{V,\boldsymbol{\eta}^\tau}) = m(P) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^{2^d} [(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^{16d^2} (1+h(P))}{\delta(\boldsymbol{\zeta})^{\min\{1/(96(d+1)k^2) - 2\epsilon^{d-l}d, \kappa(l,k^2)/4\}}} \right).$$

We shrink ϵ a final time to achieve $1/(96(d+1)k^2) - 2\epsilon^{d-l}d > 1/(100(d+1)k^2)$. The theorem follows on combining this asymptotic estimate with (8-10), when $\kappa = \kappa(d,k)$ is small in terms of $\kappa(l,k^2)$, d, and k. □

To prove Theorem 1.1 we can multiply P by any monomial, so we may assume that it is a polynomial. Thus the theorem is a direct consequence of the following more precise corollary one taking $G = \Gamma_N$ which has conductor 1.

Corollary 8.9. *Let $K \subset \mathbb{C}$ be a number field and suppose $P \in K[X_1, \dots, X_d] \backslash \{0\}$ is essentially atoral and has at most k nonzero terms for an integer $k \ge 2$. There exists $\kappa = \kappa(d,k) > 0$ with the following property. Suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order N and suppose G is a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ and $\delta(\boldsymbol{\zeta})$ is large in terms of d, P, $[K:\mathbb{Q}]$, f_G, and $[(\mathbb{Z}/N\mathbb{Z})^\times : G]$. Then $P(\boldsymbol{\zeta}^\sigma) \ne 0$ for all $\sigma \in G$ and*

$$\frac{1}{\#G} \sum_{\sigma \in G} \log|P(\boldsymbol{\zeta}^\sigma)| = m(P) + O_{d,k}\left(\frac{[K:\mathbb{Q}]^{2^d} [(\mathbb{Z}/N\mathbb{Z})^\times : G]^2 f_G \deg(P)^{16d^2} (1+h(P))}{\delta(\boldsymbol{\zeta})^{\kappa}} \right).$$

Proof. By Lemma 8.1 there is $c \ge 1$, depending only on P, such that $(P, \boldsymbol{\zeta})$ is c-admissible for all $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ of finite order with $\delta(\boldsymbol{\zeta}) \ge c$.

Suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has finite order and $P(\boldsymbol{\zeta}) = 0$. By the Manin–Mumford conjecture, $\delta(\boldsymbol{\zeta})$ is bounded in terms of d and P only. Hence for $\delta(\boldsymbol{\zeta})$ sufficiently large in terms of these quantities we have $P(\boldsymbol{\zeta}) \neq 0$. The same also holds with $\boldsymbol{\zeta}$ replaced by a Galois conjugate as $\delta(\cdot)$ is Galois invariant. Our corollary now follows from Theorem 8.8. □

Proof of Corollary 1.4. We may assume that $K/\mathbb{Q}$ is Galois and, after multiplying with a suitable monomial, that P is a polynomial. Our hypothesis implies that P is not a monomial. The product P' for $\tau(P)$ as τ ranges over $\mathrm{Gal}(K/\mathbb{Q})$ has rational coefficients. The coefficients are even integers as the coefficients of P lie in $\mathbb{Z}_K$.

The Mahler measure of any nonzero, integral polynomial is nonnegative. By a theorem attributed to Boyd [1981], Lawton [1977], and Smyth [1981], the fact that the zero set of P in $\mathbb{G}_m^d$ has an irreducible component not equal to the translate of an algebraic subgroup by a point of finite order implies $m(P') > 0$.

Suppose $\boldsymbol{\zeta} \in \mathbb{G}_m^d$ has order N. Take for G the subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ associated to $\mathrm{Gal}(\mathbb{Q}(\boldsymbol{\zeta})/K \cap \mathbb{Q}(\boldsymbol{\zeta}))$. Then $[(\mathbb{Z}/N\mathbb{Z})^\times : G] \leq [K : \mathbb{Q}]$. As $\boldsymbol{\zeta}$ varies, there are only finitely many possibilities for the number field $K \cap \mathbb{Q}(\boldsymbol{\zeta})$, being a subfield of the field K. So there is a fixed $n \in \mathbb{N}$, independent of $\boldsymbol{\zeta}$, such that $K \cap \mathbb{Q}(\boldsymbol{\zeta})$ is contained in the number field generated by a root of unity of order n. So f_G is bounded from above solely in terms of K. For any $\tau \in \mathrm{Gal}(K/\mathbb{Q})$ choose an extension $\tilde{\tau} \in \mathrm{Gal}(K(\boldsymbol{\zeta})/\mathbb{Q})$. We apply Corollary 8.9 to the polynomial $\tau(P)$ which is essentially atoral by hypothesis. If $\delta(\boldsymbol{\zeta}^{\tilde{\tau}}) = \delta(\boldsymbol{\zeta})$ is large enough in terms of the fixed data, then

$$\frac{1}{\#G} \sum_{\sigma \in G} \log|\tau(P)(\boldsymbol{\zeta}^{\tilde{\tau}\sigma})| = m(\tau(P)) + o(1)$$

as $\delta(\boldsymbol{\zeta}) \to \infty$, here and below the implied constant is independent of $\boldsymbol{\zeta}$.

The average of the left-hand side over $\tau \in \mathrm{Gal}(K/\mathbb{Q})$ equals the left-hand side in

$$\frac{1}{[K(\boldsymbol{\zeta}) : \mathbb{Q}]} \sum_{\sigma : K(\zeta) \to \mathbb{C}} \log|\sigma(P(\boldsymbol{\zeta}))| = \frac{1}{[K : \mathbb{Q}]} \sum_{\tau \in \mathrm{Gal}(K/\mathbb{Q})} m(\tau(P)) + o(1).$$

As the Mahler measure is additive, the average on the right-hand side is $m(P')/[K : \mathbb{Q}] > 0$. But the left-hand side vanishes if $P(\boldsymbol{\zeta})$ is an algebraic unit. In this case, we see that $\delta(\boldsymbol{\zeta})$ is bounded from above. □

Appendix A. A theorem of Lawton re-revisited

The following theorem makes explicit a result of Lawton [1983]. It is a more precise version of a result of Habegger [2018] which is unfortunately insufficient for our purposes. We closely follow the proof presented in [Habegger 2018] which itself is based on Lawton's approach [1983]. We also show how to correct an inaccuracy in the proof of [Habegger 2018, Lemma A.4(i)].

Recall the definition of $\rho(\cdot)$ in (6-1) where $d \geq 1$ is an integer.

Theorem A.1. *Suppose $P \in \mathbb{C}[X_1, \dots, X_d]\backslash\{0\}$ has at most k nonzero terms for an integer $k \geq 2$. For $a = (a_1, \dots, a_d) \in \mathbb{Z}^d\backslash\{0\}$ with $\rho(a) > \deg P$ we have*

$$m(P(X^{a_1}, \dots, X^{a_d})) = m(P) + O_{d,k}\left(\frac{\deg(P)^{16d^2}}{\rho(a)^{1/(16(k-1))}}\right) \tag{A-1}$$

where the implicit constant depends only on d and k.

In the univariate case $d = 1$ we have $\rho(a) = \infty$ for all $a \in \mathbb{Z}\backslash\{0\}$ by definition. Then we should interpret (A-1) as stating $m(P(X^a)) = m(P)$. This identity is an easy consequence of (4-1). So throughout this subsection we assume $d \geq 2$.

We did not strive to obtain the best-possible exponent in $\rho(a)^{1/(16(k-1))}$ that our method can produce.

We must assume $\rho(a) > \deg P$ to avoid interaction of coefficients in $P(X^{a_1}, \dots, X^{a_d})$. Indeed, take for example $P = X_1(X_2 - 1 + \epsilon)$ with $\epsilon \in (0, 1)$ small and $a = (1, 0)$. Then $P(X, 1) = X\epsilon$ whose Mahler measure is $\log \epsilon$. On the other hand $m(P) = m(X_2 - 1 + \epsilon) = \log \max\{1, |1 - \epsilon|\} = 0$ by Jensen's formula. The difference

$$m(P(X, 1)) - m(P) = \log \epsilon$$

is unbounded as $\epsilon \to 0$. This does not contradict our theorem as $\rho(a) = 1$.

The Lebesgue measure on $\mathbb{R}^d$ is denoted by $\mathrm{vol}(\cdot)$. For $P \in \mathbb{C}[X_1^{\pm 1}, \dots, X_d^{\pm 1}]$ and $r > 0$ we define

$$S(P, r) = \{x \in [0, 1)^d : |P(\boldsymbol{e}(x))| < r\} \tag{A-2}$$

where $\boldsymbol{e}$ is as in (1-3).

Dobrowolski extended Lawton's Theorem 1 [1983] to polynomials that are not necessarily monic.

Theorem A.2 [Dobrowolski 2017, Theorem 1.1]. *Suppose $P \in \mathbb{C}[X]\backslash\{0\}$ has at most k nonzero terms for an integer $k \geq 2$. Then $\mathrm{vol}(S(P, r)) \ll_k \min\{1, r/|P|\}^{1/(k-1)}$ for all $r > 0$.*

Dobrowolski requires that P as at least 2 nonzero terms. But it is convenient to allow P to have a single term, as above. It is also convenient to apply the estimate in the case $P = 0$, we then interpret the minimum to be 1.

Until the end of this appendix and if not stated otherwise we assume that $P \in \mathbb{C}[X_1, \dots, X_d]\backslash\mathbb{C}$ has at most k nonzero terms for an integer $k \geq 2$ and $|P| = 1$.

Lemma A.3. (i) *If $r > 0$ then $\mathrm{vol}(S(P, r)) \ll_{d,k} r^{1/(2(k-1))}$.*

(ii) *We have $\int_{[0,1)^d} \left|\log|P(\boldsymbol{e}(x))|\right|^2 dx \ll_{d,k} 1$.*

Proof. To ease notation we drop d, k in the subscript $\ll_{d,k}$.

Because of the trivial bound $\mathrm{vol}(S(P, r)) \leq 1$ we may assume $r \leq 1$.

The case $d = 1$ follows from Theorem A.2. So let us now assume $d \geq 2$. We consider P as a polynomial in the unknown X_d and coefficients among $\mathbb{C}[X_1, \dots, X_{d-1}]$. We pick a coefficient P_i with maximal norm, i.e., P has a term $P_i X_d^i$ such that $P_i \in \mathbb{C}[X_1, \dots, X_{d-1}]$ and $|P_i| = |P| = 1$.

For $x' \in \mathbb{R}^{d-1}$ we let $P_{\boldsymbol{e}(x')}$ denote $P(\boldsymbol{e}(x'), X) \in \mathbb{C}[X]$. Recall that

$$S(P, r) = \{(x', t) \in [0, 1)^{d-1} \times [0, 1) : |P_{\boldsymbol{e}(x')}(\boldsymbol{e}(t))| < r\}.$$

We splice the hypercube and apply Fubini's theorem to find

$$\operatorname{vol}(S(P, r)) = \int_{[0,1)^{d-1}} \operatorname{vol}(S(P_{\boldsymbol{e}(x')}, r))\, dx'.$$

The measure zero set of $x' \in [0, 1)^{d-1}$ with $P_{\boldsymbol{e}(x')} = 0$ is harmless. By Theorem A.2 we find

$$\operatorname{vol}(S(P, r)) \ll \int_{[0,1)^{d-1}} \min\left\{1, \frac{r}{|P_{\boldsymbol{e}(x')}|}\right\}^{1/(k-1)} dx'.$$

The coefficient of X^i in $P_{\boldsymbol{e}(x')}$ is $P_i(\boldsymbol{e}(x'))$. So $|P_{\boldsymbol{e}(x')}| \geq |P_i(\boldsymbol{e}(x'))|$ and

$$\operatorname{vol}(S(P, r)) \ll \int_{[0,1)^{d-1}} \min\left\{1, \frac{r}{|P_i(\boldsymbol{e}(x'))|}\right\}^{1/(k-1)} dx' = I_1 + r^{1/(k-1)} I_2 \tag{A-3}$$

where

$$I_1 = \int_{|P_i(\boldsymbol{e}(x'))|<r} dx' \quad \text{and} \quad I_2 = \int_{|P_i(\boldsymbol{e}(x'))|\geq r} \frac{dx'}{|P_i(\boldsymbol{e}(x'))|^{1/(k-1)}};$$

both integrals are over subsets of $[0, 1)^{d-1}$. We will bound I_1 and I_2 from above.

We have $I_1 = \operatorname{vol}(S(P_i, r))$. This lemma applied by induction to P_i, a polynomial in $d-1$ variables with at most k nonzero terms and $|P_i| = 1$, yields

$$I_1 \ll r^{1/(2(k-1))}. \tag{A-4}$$

To bound I_2 we consider real numbers $r = r_0 < r_1 < \cdots < r_{N+1} = k+1$, with $r_{n+1} \leq r_n + \delta$ where $\delta \in (0, 1]$ is a small parameter. We split the domain of integration up into measurable parts

$$\Sigma_n = \{x' \in [0, 1)^{d-1} : r_n \leq |P_i(\boldsymbol{e}(x'))| < r_{n+1}\} \quad \text{for } n \in \{0, \ldots, N\}.$$

Observe that $|P_i(\boldsymbol{e}(x'))| \leq k < r_{N+1}$ for all x'. Thus

$$I_2 = \sum_{n=0}^{N} \int_{\Sigma_n} \frac{dx'}{|P_i(\boldsymbol{e}(x'))|^{1/(k-1)}} \leq \sum_{n=0}^{N} \frac{\operatorname{vol}(\Sigma_n)}{r_n^{1/(k-1)}} = \sum_{n=0}^{N} a_n b_n \tag{A-5}$$

where $a_n = r_n^{-1/(k-1)}$ and $b_n = \operatorname{vol}(\Sigma_n)$.

As the Σ_n are pairwise disjoint, the partial sums satisfy

$$B_n = \sum_{l=0}^{n} b_l = \operatorname{vol}\left(\bigcup_{l=0}^{n} \Sigma_l\right) \leq \operatorname{vol}(\{x' \in [0, 1)^{d-1} : |P_i(\boldsymbol{e}(x'))| < r_{n+1}\}) = \operatorname{vol}(S(P_i, r_{n+1})).$$

In particular, we have the trivial bound $B_n \leq 1$. As in the bound for I_1 we apply this lemma by induction to P_i and find

$$0 \leq B_n \leq \operatorname{vol}(S(P_i, r_{n+1})) \ll r_{n+1}^{1/(2(k-1))}. \tag{A-6}$$

Summation by parts implies

$$I_2 \le \sum_{n=0}^{N} a_n b_n = a_N B_N - \sum_{n=0}^{N-1} B_n(a_{n+1} - a_n) \le 1 + \sum_{n=0}^{N-1} B_n(a_n - a_{n+1});$$

we used $a_N = r_N^{-1/(k-1)} \le 1$ as $r_N \ge r_{N+1} - \delta \ge 1$ and $B_N \le 1$. By (A-6) and the definition of a_n we find

$$I_2 \ll 1 + \sum_{n=0}^{N-1} r_{n+1}^{1/(2(k-1))}(r_n^{-1/(k-1)} - r_{n+1}^{-1/(k-1)}).$$

We use the mean value theorem to bound

$$r_n^{-1/(k-1)} - r_{n+1}^{-1/(k-1)} \ll r_n^{-1/(k-1)-1}(r_{n+1} - r_n) \ll r_{n+1}^{-1/(k-1)-1}(r_{n+1} - r_n);$$

for the second bound we assume, as we may, that $\delta \le r$ and so $r_{n+1} \le r_n + \delta \le 2r_n$. Thus $I_2 \ll 1 + \int_r^{k+1} t^{-1/(2(k-1))-1} dt \ll r^{-1/(2(k-1))}$.

This bound together with (A-4) implies $I_1 + r^{1/(k-1)} I_2 \ll r^{1/(2(k-1))}$. Therefore, $\mathrm{vol}(S(P,r)) \ll r^{1/(2(k-1))}$ by (A-3), completing the induction step and the proof of (i).

We define $p_n(x) = \min\{n, |\log|P(\boldsymbol{e}(x))||^2\} \ge 0$ where $n \ge 0$ is an integer. We must find an upper bound for the nondecreasing sequence $I_n = \int_{[0,1)^d} p_n(x)\,dx$. Observe that $|P(\boldsymbol{e}(x))| \le k|P| = k$, so if $n \ge (\log k)^2$, then $|P(\boldsymbol{e}(x))| \le e^{\sqrt{n}}$. We fix m to be the least integer with $m \ge 1 + (\log k)^2$, so $m \ge 2$. Say $n \ge m$. Then p_n equals n on $S(P, e^{-\sqrt{n}})$ and it equals p_{n+1} outside this set. Thus

$$I_{n+1} - I_n = \int_{S(P,e^{-\sqrt{n}})} (p_{n+1}(x) - p_n(x))\,dx \le \mathrm{vol}(S(P, e^{-\sqrt{n}})) \ll e^{-\lambda\sqrt{n}}$$

from part (i), here $\lambda = 1/(2(k-1))$. A telescoping sum trick shows

$$I_n - I_m \ll \sum_{l \ge m} e^{-\lambda\sqrt{l}} \ll \int_{m-1}^{\infty} e^{-\lambda\sqrt{l}}\,dl \ll 1.$$

The initial term satisfies $I_m \le m \ll 1$ as m depends only on k, this completes the proof. □

A more careful analysis should lead to $\mathrm{vol}(S(P,r)) \ll_{d,k} (1 + |\log r|)^{d-1} r^{1/(k-1)}$ for all $r > 0$ in part (i) of Lemma A.3. But this improvement has little effect on the main results of the current work.

Brunault, Guilloux, Mehrabdollahi, and Pengo pointed out that the argument for second-named author's [Habegger 2018, Lemma A.4(i)] leads (for $k \ge 2$) to an estimate $O(y^{f(n)/(2(k-1))})$ where $f(n)$ depends on the number of variables n, as opposed to the claimed bound $O(y^{1/(2(k-1))})$. However, the claimed bound holds true by Lemma A.3(i). Alternatively and in the proof of Lemma A.3(i) one can replace Dobrowolski's Theorem 1.1 [2017] by Lawton's Theorem 1 [1983] which is sufficient for the applications in [Habegger 2018].

Lemma A.4. *If $r > 0$ then*

$$\int_{S(P,r)} \big|\log|P(\boldsymbol{e}(x))|\big|\,dx \ll_{d,k} r^{1/(4(k-1))}.$$

Proof. As $|P(\boldsymbol{e}(x))| \le |P|k \le k$ for all $x \in [0,1)^d$ we may assume $r \le 1$ by the Cauchy–Schwarz inequality and Lemma A.3(ii).

With $\Sigma = S(P, r)$ we find

$$0 \le -\int_\Sigma \log|P(\boldsymbol{e}(x))|\,dx = -\sum_{n=0}^{\infty} \int_{r/2^{n+1} \le |P(\boldsymbol{e}(x))| < r/2^n} \log|P(\boldsymbol{e}(x))|\,dx \le \sum_{n=0}^{\infty} \log\left(\frac{2^{n+1}}{r}\right) \mathrm{vol}(S(P, r/2^n)).$$

Let $\lambda = 1/(2(k-1)) \le \frac{1}{2}$. We use Lemma A.3(i) to bound $\mathrm{vol}(S(P, r/2^n)) \ll_{d,k} (r/2^n)^\lambda$. Note that $\log(2t) \ll_k t^{\lambda/2}$ on $t \in [1, \infty)$. We take $t = 2^n/r \ge 1$ and conclude

$$-\int_\Sigma \log|P(\boldsymbol{e}(x))|\,dx \ll_{d,k} \sum_{n=0}^{\infty} \left(\frac{r}{2^n}\right)^{\lambda/2} \ll_{d,k} r^{\lambda/2}. \qquad \square$$

Boyd [1998] proved that the Mahler measure is continuous on the nonzero polynomials of fixed degree. Here we show that the Mahler measure is Hölder continuous away from 0. For the next lemma we momentarily drop our usual assumptions on P.

Lemma A.5. *Suppose $P, Q \in \mathbb{C}[X_1, \dots, X_d]\backslash\{0\}$ such that P and Q both have at most k nonzero terms for an integer $k \ge 2$. If $\delta = |P-Q|/|Q| \le \frac{1}{2}$, then*

$$m(P) \le m(Q) + C(d,k)\delta^{1/(8(k-1))}$$

where $C(d,k) > 0$ is effective and depends only on d and k.

Proof. It suffices to prove the lemma when $|Q| = 1$; indeed, just replace P and Q by $P/|Q|$ and $Q/|Q|$, respectively, to reduce to this case.

Suppose for the moment that $x \in \mathbb{R}^d$ with $P(\boldsymbol{e}(x))Q(\boldsymbol{e}(x)) \ne 0$. Then $|P(\boldsymbol{e}(x)) - Q(\boldsymbol{e}(x))| \le 2k|P-Q| = 2k\delta$ and so

$$\log\left|\frac{P(\boldsymbol{e}(x))}{Q(\boldsymbol{e}(x))}\right| \le \left|\frac{P(\boldsymbol{e}(x))}{Q(\boldsymbol{e}(x))}\right| - 1 \le 2k\frac{\delta}{|Q(\boldsymbol{e}(x))|} \tag{A-7}$$

where the first inequality used $\log t \le t - 1$ for all $t > 0$.

The difference of Mahler measures $m(P) - m(Q)$ can be written as

$$\int_{[0,1)^d\backslash\Sigma} (\log|P(\boldsymbol{e}(x))| - \log|Q(\boldsymbol{e}(x))|)\,dx + \int_\Sigma (\log|P(\boldsymbol{e}(x))| - \log|Q(\boldsymbol{e}(x))|)\,dx$$

with $\Sigma = S(Q, \delta^{1/2})$.

The first integral is at most $2k\delta^{1/2}$ by (A-7). We proceed by bounding the second integral I from above. First, we note that $|P(\boldsymbol{e}(x))| \le k|P| \le 3k/2$ as $|P-Q| \le \delta \le \frac{1}{2}$ and thus $|P| \le \frac{3}{2}$. So

$$I \le \log(3k/2)\mathrm{vol}(\Sigma) - \int_\Sigma \log|Q(\boldsymbol{e}(x))|\,dx \le \log(3k/2)\mathrm{vol}(\Sigma) + c\delta^{1/(8(k-1))}$$

where we applied Lemma A.4 to Q and $\delta^{1/2}$, the case Q constant being trivial; here $c = c(d,k) > 0$ depends only on d and k. Finally, Lemma A.3(i) yields $\mathrm{vol}(\Sigma) = \mathrm{vol}(S(Q,\delta^{1/2})) \ll_{d,k} \delta^{1/(4(k-1))}$ and the lemma follows as $\delta \le 1$. □

Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $b \in \mathbb{N}_0$ let $C^b(\mathbb{R}^d)$ denote the set of real valued functions on $\mathbb{R}^d$ whose derivatives exist and are continuous up to and including order b. For a multiindex $i = (i_1, \dots, i_d) \in \mathbb{N}_0^d$ we set $\ell(i) = i_1 + \cdots + i_d$. If $g \in C^b(\mathbb{R}^d)$ and $\ell(i) \le b$, we set $\partial^i g = (\partial/\partial x_1)^{i_1} \cdots (\partial/\partial x_d)^{i_d} g \in C^0(\mathbb{R}^d)$ and

$$|g|_{C^b} = \max_{\substack{i \in \mathbb{N}_0^d \\ \ell(i) \le b}} \sup_{x \in \mathbb{R}^d} |\partial^i g(x)| \in \mathbb{R} \cup \{\infty\}.$$

We recall here the construction of $f_r \in C^b(\mathbb{R}^d)$ as in [Habegger 2018] where $r \in \left(0, \frac{1}{2}\right]$ is a parameter. This function equals $\log|P(\,\cdot\,))|$ away from the singularity, i.e., the locus where $P(\boldsymbol{e}(\,\cdot\,))$ vanishes.

We fix the antiderivative ϕ of $x^b(1-x)^b$ on $[0,1]$ with $\phi(0) = 0$ and multiply it with a positive number to ensure $\phi(1) = 1$. Then we extend it by 0 on $x < 0$ and by 1 for $x > 1$ to obtain a nondecreasing step function $\phi \in C^b(\mathbb{R})$ whose derivative ϕ' has support $[0,1]$. Finally, we rescale and define $\phi_r(x) = \phi(((2/r)^2 x - 1)/3)$. So ϕ_r is a nondecreasing function which vanishes on $(-\infty, (r/2)^2]$, equals 1 on $[r^2, \infty)$, and satisfies

$$\left|\frac{d^i \phi_r}{dx^i}\right|_{C^0} \ll_b r^{-2i} \quad \text{for all } 0 \le i \le b, \text{ hence} \quad |\phi_r|_{C^b} \ll_b r^{-2b}.$$

The function ϕ_r takes values in $[0,1]$. Moreover, we define

$$\psi_r(x) = \begin{cases} \frac{1}{2}\phi_r(x)\log x, & x > 0, \\ 0, & x \le 0. \end{cases}.$$

Then ψ_r vanishes on $(-\infty, (r/2)^2]$, coincides with $\frac{1}{2}\log x$ on $[r^2, \infty)$, and satisfies

$$|\psi_r|_{C^b} \ll_b r^{-2b} |\log r|. \tag{A-8}$$

We consider $g : x \mapsto |P(\boldsymbol{e}(x))|^2$, then

$$|g|_{C^b} \ll_{k,b} (\deg P)^b; \tag{A-9}$$

recall that $|P| = 1$. Next we compose $f_r = \psi_r \circ g \in C^b(\mathbb{R}^d)$, so for $x \in \mathbb{R}^d$ we have

$$f_r(x) = \begin{cases} 0, & \text{if } |P(\boldsymbol{e}(x))| \le r/2, \\ \log|P(\boldsymbol{e}(x))|, & \text{if } |P(\boldsymbol{e}(x))| \ge r. \end{cases}.$$

By [Habegger 2018, Lemma A.5], which follows from the chain rule, together with (A-8) and (A-9) we find

$$|f_r|_{C^b} \ll_{k,b} r^{-2b} |\log r| (\deg P)^{b^2}. \tag{A-10}$$

For the following lemmas we suppose $b \ge d+1$. As above we have $r \in \left(0, \frac{1}{2}\right]$.

Lemma A.6. *Suppose* $a \in \mathbb{Z}^d \backslash \{0\}$, *then*

$$\int_0^1 f_r(as)\,ds = \int_{[0,1)^d} f_r(x)\,dx + O_{d,k,b}\left(\frac{|\log r|}{r^{2b}}\frac{(\deg P)^{b^2}}{\rho(a)^{b-d}}\right).$$

We follow and adapt the proof of [Habegger 2018, Lemma A.6].

Proof. For $m \in \mathbb{Z}^d$ let $\widehat{f_r}(m)$ denote the Fourier coefficient of f_r. By [Grafakos 2014, Theorem 3.2.9(a)] with derivative up to order b and using $|\widehat{\partial^i f_r}(m)| \le |\partial^i f_r|_{C^0} \le |f_r|_{C^b}$ where $\ell(i) = b$ we conclude $|\widehat{f_r}(m)| \ll_{d,k,b} |f_r|_{C^b}|m|^{-b}$ if $m \neq 0$. So $|\widehat{f_r}(m)| \ll_{d,k,b} r^{-2b}|\log r|(\deg P)^{b^2}|m|^{-b}$ for all $m \in \mathbb{Z}^d\backslash\{0\}$ by (A-10). Then

$$\sum_{|m|\ge\rho(a)} |\widehat{f_r}(m)| \ll_{d,k,b} \frac{|\log r|}{r^{2b}}(\deg P)^{b^2} \sum_{|m|\ge\rho(a)} \frac{1}{|m|^b} \ll_{d,k,b} \frac{|\log r|}{r^{2b}}\frac{(\deg P)^{b^2}}{\rho(a)^{b-d}} \tag{A-11}$$

as $b \ge d+1$. In particular, the Fourier coefficients of f_r are absolutely summable and the Fourier series converges absolutely and uniformly to f_r, see [Grafakos 2014, Proposition 3.1.14]. Hence

$$\int_0^1 f_r(as)\,ds = \sum_{m\in\mathbb{Z}^d} \int_0^1 \widehat{f_r}(m)e^{2\pi\sqrt{-1}\langle a,m\rangle s}\,ds = \int_{[0,1)^d} f_r(x)\,dx + \sum_{\substack{m\in\mathbb{Z}^d\backslash\{0\}\\ \langle a,m\rangle=0}} \widehat{f_r}(m).$$

The lemma follows from (A-11) as only those m with $|m| \ge \rho(a)$ contribute to the final sum. □

Lemma A.7. *Suppose* $a \in \mathbb{Z}^d\backslash\{0\}$ *such that* $\rho(a) > \deg P$. *For all* $s \in [0,1)$, *up to finitely many exceptions, we have* $|P(\boldsymbol{e}(as))| \neq 0$ *and*

$$\int_0^1 \log|P(\boldsymbol{e}(as))|\,ds = \int_0^1 f_r(as)\,ds + O_k(r^{1/(k-1)}|\log r|).$$

We follow and adapt the proof of [Habegger 2018, Lemma A.7].

Proof. Say $a = (a_1, \dots, a_d)$ with $\rho(a) > \deg P$. Then the coefficients of the univariate Laurent polynomial $Q = P(X^{a_1}, \dots, X^{a_d})$ are precisely the coefficients of P. Hence $|Q| = |P| = 1$ and Q has at most k nonzero terms.

The first claim follows as $P(\boldsymbol{e}(as)) = Q(\boldsymbol{e}(s))$ for all $s \in \mathbb{R}$ and since $Q \neq 0$.

To prove the second claim we note that the difference of the two integrals equals

$$\int_{S(Q,r)} (\log|Q(\boldsymbol{e}(s))| - f_r(as))\,ds$$

with $S(Q,r)$ as in (A-2). Note that $\int_{S(Q,r)} \log|Q(\boldsymbol{e}(s))|\,ds \le 0$ as $r \le 1$. Recall Theorem A.2 which yields $\mathrm{vol}(S(Q,r)) \ll_k r^{1/(k-1)}$. As in the proof of [Lawton 1983, Lemma 4], see also [Schinzel 2000, Theorem 7, Appendix G], we find

$$\int_{S(Q,r)} \log|Q(\boldsymbol{e}(s))|\,ds \ge -Cr^{1/(k-1)}|\log r|,$$

where $C > 0$ depends only on k. Finally, by the definition of f_r we find $\log(r/2) \le f_r(as) \le 0$ if $|Q(\boldsymbol{e}(s))| < r$. Thus $\int_{S(Q,r)} f_r(as)\,ds$ is also $O_k(r^{1/(k-1)}|\log r|)$. □

Lemma A.8. *We have*

$$\left|\int_{[0,1)^d} (f_r(x) - \log|P(\boldsymbol{e}(x))|)\,dx\right| \ll_{d,k} r^{1/(4(k-1))}.$$

We follow and adapt the proof of [Habegger 2018, Lemma A.8].

Proof. We have

$$\begin{aligned}\left|\int_{[0,1)^d} (f_r(x)-\log|P(\boldsymbol{e}(x))|)\,dx\right| &= \left|\int_{[0,1)^d} (\phi_r(|P(\boldsymbol{e}(x))|^2)-1)\log|P(\boldsymbol{e}(x))|\,dx\right| \\ &\le \int_{[0,1)^d} \left|\phi_r(|P(\boldsymbol{e}(x))|^2)-1\right|\left|\log|P(\boldsymbol{e}(x))|\right|dx \\ &\le \left(\int_{[0,1)^d} \left|\phi_r(|P(\boldsymbol{e}(x))|^2)-1\right|^2 dx\right)^{1/2}\left(\int_{[0,1)^d} \left|\log|P(\boldsymbol{e}(x))|\right|^2 dx\right)^{1/2}\end{aligned}$$

by the definition of f_r and where we used the Cauchy–Schwarz inequality in the last step. The second integral on the final line is $\ll_{d,k} 1$ by Lemma A.3(ii). The first integral is

$$\int_{S(P,r)} \left|\phi_r(|P(\boldsymbol{e}(x))|^2) - 1\right|^2 dx \le \mathrm{vol}(S(P,r)) \ll_{d,k} r^{1/(2(k-1))}$$

by Lemma A.3(i). We take the square root to complete the proof. □

Proof of Theorem A.1. As stated below Theorem A.1 we may assume $d \ge 2$. We may also assume that P is nonconstant. As we have seen in the proof of Lemma A.7, the condition $\rho(a) > \deg P$ guarantees $P(X^{a_1}, \dots, X^{a_d}) \ne 0$. Moreover, replacing P by $P/|P|$ leaves $m(P(X^{a_1}, \dots, X^{a_d})) - m(P)$ invariant. So it suffices to prove the theorem if $|P| = 1$.

We fix the parameters $b = 4d \ge d+1$ and $r = \rho(a)^{-1/4}/2 \le \frac{1}{2}$.

We write $|m(P(X^{a_1}, \dots, X^{a_d})) - m(P)|$ as $\int_0^1 \log|P(\boldsymbol{e}(as))|\,ds - \int_{[0,1)^d} \log|P(\boldsymbol{e}(x))|\,dx$ and find that it is at most

$$\left|\int_0^1 f_r(as)\,ds - \int_{[0,1)^d} f_r(x)\,dx\right| + \left|\int_0^1 (\log|P(\boldsymbol{e}(as))| - f_r(as))\,ds\right| + \left|\int_{[0,1)^d} (f_r(x)-\log|P(\boldsymbol{e}(x))|)\,dx\right|.$$

Then by Lemmas A.6, A.7, and A.8 this sum is

$$\ll_{d,k} \frac{|\log r|}{r^{2b}} \frac{(\deg P)^{b^2}}{\rho(a)^{b-d}} + r^{1/(k-1)}|\log r| + r^{1/(4(k-1))}.$$

By our choice of r and $\rho(a) \ge 2$, the sum is

$$\ll_{d,k} \frac{\log \rho(a)}{\rho(a)^{b-d-b/2}} (\deg P)^{b^2} + \frac{\log \rho(a)}{\rho(a)^{1/(4(k-1))}} + \frac{1}{\rho(a)^{1/(16(k-1))}}.$$

Finally, as $b = 4d$ the sum is

$$\ll_{d,k} (\deg P)^{16d^2} \frac{\log \rho(a)}{\rho(a)^d} + \frac{\log \rho(a)}{\rho(a)^{1/(4(k-1))}} + \frac{1}{\rho(a)^{1/(16(k-1))}}. \qquad \square$$

Appendix B. Recovering the theorem of Lind, Schmidt, and Verbitskiy

In this appendix we recover from our work a variant of Lind, Schmidt, and Verbitskiy's Theorem 1.1 [2013]. This variant is stated in the introduction as Theorem 1.2. Let $d \in \mathbb{N}$. For a finite subgroup $G \subset \mathbb{G}_m^d$, recall that we defined $\delta(G)$ in (1-5).

Lemma B.1. *Let G be a finite subgroup of $\mathbb{G}_m^d$. If $a \in \mathbb{Z}^d \backslash \{0\}$, then*

$$\frac{1}{\#G} \#\{\zeta \in G : \zeta^a = 1\} \le \frac{|a|}{\delta(G)}.$$

Proof. We will detect $\zeta^a = 1$ using the character $\chi(\zeta) = \zeta^a$ of G. The image $\chi(G)$ is a cyclic subgroup of $\mathbb{C}^\times$ of order N, say. For $\zeta \in G$, the sum $\sum_{k=0}^{N-1} \chi(\zeta)^k = 0$ equals N if $\zeta^a = 1$ and vanishes otherwise. The number of solutions $\zeta \in G$ of $\zeta^a = 1$ is thus

$$\sum_{\zeta \in G} \frac{1}{N} \sum_{k=0}^{N-1} \chi(\zeta^k) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{\zeta \in G} \chi(\zeta)^k = \frac{1}{N} \sum_{k=0}^{N-1} \frac{\#G}{N} \sum_{\xi \in \chi(G)} \xi^k = \frac{\#G}{N}.$$

We conclude the proof as $\zeta^{aN} = \chi(\zeta)^N = 1$ for all $\zeta \in G$ and hence $N \ge \delta(G)/|a|$. $\square$

Lemma B.2. *Let G be a finite subgroup of $\mathbb{G}_m^d$:*

(i) *If $T \ge 1$, then*

$$\frac{1}{\#G} \#\{\zeta \in G : \delta(\zeta) \le T\} \le \frac{3^d T^{d+1}}{\delta(G)}.$$

(ii) *If $\kappa > 0$, then*

$$\frac{1}{\#G} \sum_{\zeta \in G} \delta(\zeta)^{-\kappa} \le \frac{4^d}{\delta(G)^{\kappa/(d+1+\kappa)}}.$$

Proof. Any $\zeta \in G$ with $\delta(\zeta) \le T$ satisfies $\zeta^a = 1$ for some $a \in \mathbb{Z}^d \backslash \{0\}$ and $|a| \le T$. The number of such a is at most $(2T+1)^d \le 3^d T^d$ and each a leads to at most $|a|\#G/\delta(G) \le T\#G/\delta(G)$ different ζ by Lemma B.1. This implies (i).

For the second assertion we split up the elements in G into those with $\delta(\zeta) \le T$ and those with $\delta(\zeta) > T$; here $T \ge 1$ is a parameter to be chosen.

For the lower range, we use the trivial lower bound $\delta(\zeta) \ge 1$ and part (i) to obtain

$$\frac{1}{\#G} \sum_{\substack{\zeta \in G \\ \delta(\zeta) \le T}} \delta(\zeta)^{-\kappa} \le \frac{3^d T^{d+1}}{\delta(G)}.$$

For the higher range, we have

$$\frac{1}{\#G} \sum_{\substack{\boldsymbol{\zeta} \in G \\ \delta(\boldsymbol{\zeta}) > T}} \delta(\boldsymbol{\zeta})^{-\kappa} \leq \frac{1}{T^{\kappa}}.$$

The lemma follows by taking the sum of these two bounds with $T = \delta(G)^{1/(d+1+\kappa)}$. □

Proof of Theorem 1.2. Without loss of generality we can assume that P is a polynomial.

Any finite subgroup of $\mathbb{G}_m^d$ is defined over $\mathbb{Q}$, i.e., it is mapped to itself under the action of the absolute Galois group of $\mathbb{Q}$, see [Bombieri and Gubler 2006, Corollary 3.2.15]. We decompose G into a disjoint union $G_1 \cup \cdots \cup G_m$ of Galois orbits. It is useful to fix a representative $\boldsymbol{\zeta}_i \in G_i$ for each $i \in \{1, \ldots, m\}$ and define $N_i = \operatorname{ord}(\boldsymbol{\zeta}_i)$. All elements in G_i have the same order and the Galois action is the natural action of $(\mathbb{Z}/N_i\mathbb{Z})^{\times}$ on G_i. Moreover, $\#G_i = \varphi(N_i)$. Note that δ is constant on each G_i as $\delta(\boldsymbol{\zeta}^{\sigma}) = \delta(\boldsymbol{\zeta})$ for all field automorphisms σ. Moreover, $P(\boldsymbol{\zeta}_i) \neq 0$ if and only if P does not vanish at any point of G_i; indeed P has rational coefficients by hypothesis.

Let $T \geq 1$ be a parameter depending on $\delta(G)$ and large in terms of P, d which we will fix in due time. We split our average (1-6) up into those $\boldsymbol{\zeta}$ with $\delta(\boldsymbol{\zeta}) \leq T$ and those with $\delta(\boldsymbol{\zeta}) > T$.

First, we will show that the sum

$$\frac{1}{\#G} \sum_{\substack{\boldsymbol{\zeta} \in G \\ \delta(\boldsymbol{\zeta}) \leq T, P(\boldsymbol{\zeta}) \neq 0}} \log|P(\boldsymbol{\zeta})| = \frac{1}{\#G} \sum_{\substack{i=1 \\ \delta(\boldsymbol{\zeta}_i) \leq T, P(\boldsymbol{\zeta}_i) \neq 0}}^{m} \sum_{\sigma \in (\mathbb{Z}/N_i\mathbb{Z})^{\times}} \log|P(\boldsymbol{\zeta}_i^{\sigma})| \tag{B-1}$$

is negligible. Say $P(\boldsymbol{\zeta}_i) \neq 0$. Then $P(\boldsymbol{\zeta}_i)$ lies in a number field of degree $\varphi(N_i)$ over $\mathbb{Q}$. So

$$\left| \sum_{\sigma \in (\mathbb{Z}/N_i\mathbb{Z})^{\times}} \log|P(\boldsymbol{\zeta}_i^{\sigma})| \right| \leq \sum_{\sigma \in (\mathbb{Z}/N_i\mathbb{Z})^{\times}} \left| \log|P(\boldsymbol{\zeta}_i^{\sigma})| \right| \leq 2\varphi(N_i) h(P(\boldsymbol{\zeta}_i)) \ll_P \varphi(N_i)$$

where we used basic properties of the height (2-3) and $P(\boldsymbol{\zeta}_i^{\sigma}) = P(\boldsymbol{\zeta}_i)^{\sigma}$ for all $\sigma \in \operatorname{Gal}(\mathbb{Q}(\boldsymbol{\zeta}_i)/\mathbb{Q})$. So the absolute value of (B-1) is at most

$$\ll_P \frac{1}{\#G} \sum_{\substack{i=1 \\ \delta(\boldsymbol{\zeta}_i) \leq T}}^{m} \varphi(N_i) \ll_P \frac{1}{\#G} \sum_{\substack{\boldsymbol{\zeta} \in G \\ \delta(\boldsymbol{\zeta}) \leq T}} 1 \ll_{d,P} \frac{T^{d+1}}{\delta(G)}. \tag{B-2}$$

by Lemma B.2(i).

The remaining sum is

$$\frac{1}{\#G} \sum_{\substack{i=1 \\ \delta(\boldsymbol{\zeta}_i) > T}}^{m} \sum_{\sigma \in (\mathbb{Z}/N_i\mathbb{Z})^{\times}} \log|P(\boldsymbol{\zeta}_i^{\sigma})|;$$

note that $P(\boldsymbol{\zeta}_i^\sigma) \neq 0$ for T large enough by the Manin–Mumford conjecture for $\mathbb{G}_m^d$. Using Theorem 1.1 we have

$$\frac{1}{\#G}\sum_{\substack{i=1\\ \delta(\boldsymbol{\zeta}_i)>T}}^{m}\varphi(N_i)(m(P)+O_{d,P}(\delta(\boldsymbol{\zeta}_i)^{-\kappa})) = \frac{1}{\#G}\Bigg(\sum_{\boldsymbol{\zeta}\in G:\delta(\boldsymbol{\zeta})>T}1\Bigg)m(P)+O_{d,P}\Bigg(\frac{1}{\#G}\sum_{\boldsymbol{\zeta}\in G:\delta(\boldsymbol{\zeta})>T}\delta(\boldsymbol{\zeta})^{-\kappa}\Bigg)$$
$$= \Bigg(1-\frac{1}{\#G}\sum_{\boldsymbol{\zeta}\in G,\delta(\boldsymbol{\zeta})\leq T}1\Bigg)m(P)+O_{d,P}\Big(\delta(G)^{-\kappa/(d+1+\kappa)}\Big)$$

where we used Lemma B.2(ii). The remaining average in the last line is $O_d(T^{d+1}/\delta(G))$ by Lemma B.2(i).

We combine this estimate with the first bound (B-2) to conclude that the average (1-6) equals

$$m(P)+O_{d,P}(T^{d+1}\delta(G)^{-1}+\delta(G)^{-\kappa/(d+1+\kappa)})$$

The theorem follows with the choice $T = c\delta(G)^{1/(2(d+1))}$ where $c \geq 1$ is sufficiently large in terms of d and P. The exponent κ in (1-6) is $\min\{\frac{1}{2}, \kappa/(d+1+\kappa)\}$ in the notation here. □

We leave to the interested reader the task of generalizing the previous theorem to polynomials defined over an arbitrary number field.

Acknowledgments

The authors thank Pierre Le Boudec for references regarding Gauss sums, Peter Sarnak for the reference to Le's work [2014], and Shouwu Zhang for pointing out the work of Chambert-Loir and Thuillier [2009]. We also thank the referee for carefully reading this text and for providing many valuable comments that led to improvements of the text and some simplifications. The authors thank François Brunault, Antonin Guilloux, Mahya Mehrabdollahi, and Riccardo Pengo for pointing out a mistake in an earlier attempt to prove Lemma A.3(i) and for making us aware of the work of Dobrowolski [2017]. Vesselin Dimitrov gratefully acknowledges support from the European Research Council via ERC grant GeTeMo 617129. Philipp Habegger has received funding from the Swiss National Science Foundation project n° 200020_184623.

References

[Agler et al. 2006] J. Agler, J. E. McCarthy, and M. Stankus, "Toral algebraic sets and function theory on polydisks", *J. Geom. Anal.* **16**:4 (2006), 551–562. MR Zbl

[Autissier 2006] P. Autissier, "Sur une question d'équirépartition de nombres algébriques", *C. R. Math. Acad. Sci. Paris* **342**:9 (2006), 639–641. MR Zbl

[Baker et al. 2008] M. Baker, S.-i. Ih, and R. Rumely, "A finiteness property of torsion points", *Algebra Number Theory* **2**:2 (2008), 217–248. MR Zbl

[Bilu 1997] Y. Bilu, "Limit distribution of small points on algebraic tori", *Duke Math. J.* **89**:3 (1997), 465–476. MR Zbl

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Math. Monogr. **4**, Cambridge Univ. Press, 2006. MR Zbl

[Bombieri et al. 2007] E. Bombieri, D. Masser, and U. Zannier, "Anomalous subvarieties: structure theorems and applications", *Int. Math. Res. Not.* **2007**:19 (2007), art. id. rnm057. MR Zbl

[Boyd 1981] D. W. Boyd, "Kronecker's theorem and Lehmer's problem for polynomials in several variables", *J. Number Theory* **13**:1 (1981), 116–121. MR Zbl

[Boyd 1998] D. W. Boyd, "Uniform approximation to Mahler's measure in several variables", *Canad. Math. Bull.* **41**:1 (1998), 125–128. MR Zbl

[Brunault et al. 2022] F. Brunault, A. Guilloux, M. Mehrabdollahei, and R. Pengo, "Limits of Mahler measures in multiple variables", preprint, 2022. arXiv 2203.12259

[Cassels 1959] J. W. S. Cassels, *An introduction to the geometry of numbers*, Grundl. Math. Wissen. **99**, Springer, 1959. MR Zbl

[Chambert-Loir and Thuillier 2009] A. Chambert-Loir and A. Thuillier, "Mesures de Mahler et équidistribution logarithmique", *Ann. Inst. Fourier* (*Grenoble*) **59**:3 (2009), 977–1014. MR Zbl

[Dimitrov 2016] V. Dimitrov, "Convergence to the Mahler measure and the distribution of periodic points for algebraic Noetherian $\mathbb{Z}^d$-actions", preprint, 2016. arXiv 1611.04664

[Dobrowolski 2017] E. Dobrowolski, "A note on Lawton's theorem", *Canad. Math. Bull.* **60**:3 (2017), 484–489. MR Zbl

[Dobrowolski and Smyth 2017] E. Dobrowolski and C. Smyth, "Mahler measures of polynomials that are sums of a bounded number of monomials", *Int. J. Number Theory* **13**:6 (2017), 1603–1610. MR Zbl

[Dubickas 1997] A. Dubickas, "Algebraic conjugates outside the unit circle", pp. 11–21 in *New trends in probability and statistics*, *IV* (Palanga, Lithuania, 1996), edited by A. Laurinčikas et al., VSP, Utrecht, Netherlands, 1997. MR Zbl

[Dubickas 2018] A. Dubickas, "On sums of two and three roots of unity", *J. Number Theory* **192** (2018), 65–79. MR Zbl

[Duke 2007] W. Duke, "A combinatorial problem related to Mahler's measure", *Bull. Lond. Math. Soc.* **39**:5 (2007), 741–748. MR Zbl

[Grafakos 2014] L. Grafakos, *Classical Fourier analysis*, 3rd ed., Grad. Texts in Math. **249**, Springer, 2014. MR Zbl

[Grayson 1984] D. R. Grayson, "Reduction theory using semistability", *Comment. Math. Helv.* **59**:4 (1984), 600–634. MR Zbl

[Güting 1961] R. Güting, "Approximation of algebraic numbers by algebraic numbers", *Michigan Math. J.* **8** (1961), 149–159. MR Zbl

[Habegger 2018] P. Habegger, "The norm of Gaussian periods", *Q. J. Math.* **69**:1 (2018), 153–182. MR Zbl

[Harman 1998] G. Harman, *Metric number theory*, Lond. Math. Soc. Monogr. (N.S.) **18**, Oxford Univ. Press, 1998. MR Zbl

[Hlawka 1971] E. Hlawka, "Discrepancy and Riemann integration", pp. 121–129 in *Studies in pure mathematics*, edited by L. Mirsky, Academic Press, London, 1971. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ. **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Kuipers and Niederreiter 1974] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, New York, 1974. MR Zbl

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Grad. Texts in Math. **211**, Springer, 2002. MR Zbl

[Laurent 1984] M. Laurent, "Équations diophantiennes exponentielles", *Invent. Math.* **78**:2 (1984), 299–327. MR Zbl

[Laurent et al. 1995] M. Laurent, M. Mignotte, and Y. Nesterenko, "Formes linéaires en deux logarithmes et déterminants d'interpolation", *J. Number Theory* **55**:2 (1995), 285–321. MR Zbl

[Lawton 1977] W. M. Lawton, "A generalization of a theorem of Kronecker", *J. Sci. Fat. Chiangmai Unit.* **4** (1977), 15–23. Zbl

[Lawton 1983] W. M. Lawton, "A problem of Boyd concerning geometric means of polynomials", *J. Number Theory* **16**:3 (1983), 356–362. MR Zbl

[Le 2014] T. Le, "Homology torsion growth and Mahler measure", *Comment. Math. Helv.* **89**:3 (2014), 719–757. MR Zbl

[Lind et al. 2010] D. Lind, K. Schmidt, and E. Verbitskiy, "Entropy and growth rate of periodic points of algebraic $\mathbb{Z}^d$-actions", pp. 195–211 in *Dynamical numbers*: *interplay between dynamical systems and number theory* (Bonn, Germany, 2009), edited by S. Kolyada et al., Contemp. Math. **532**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Lind et al. 2013] D. Lind, K. Schmidt, and E. Verbitskiy, "Homoclinic points, atoral polynomials, and periodic points of algebraic $\mathbb{Z}^d$-actions", *Ergodic Theory Dynam. Systems* **33**:4 (2013), 1060–1081. MR Zbl

[Mahler 1964] K. Mahler, "An inequality for the discriminant of a polynomial", *Michigan Math. J.* **11** (1964), 257–262. MR Zbl

[Mignotte 1995] M. Mignotte, "On the distance between the roots of a polynomial", *Appl. Algebra Engrg. Comm. Comput.* **6**:6 (1995), 327–332. MR Zbl

[Myerson 1979] G. Myerson, "A combinatorial problem in finite fields, I", *Pacific J. Math.* **82**:1 (1979), 179–187. MR Zbl

[Myerson 1980] G. Myerson, "A combinatorial problem in finite fields, II", *Q. J. Math.* **31**:122 (1980), 219–231. MR Zbl

[Myerson 1986] G. Myerson, "Unsolved problems: how small can a sum of roots of unity be?", *Amer. Math. Monthly* **93**:6 (1986), 457–459. MR Zbl

[Rahman and Schmeisser 2002] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, Lond. Math. Soc. Monogr. (N.S.) **26**, Oxford Univ. Press, 2002. MR Zbl

[Rosser and Schoenfeld 1962] J. B. Rosser and L. Schoenfeld, "Approximate formulas for some functions of prime numbers", *Illinois J. Math.* **6** (1962), 64–94. MR Zbl

[Sarnak and Adams 1994] P. Sarnak and S. Adams, "Betti numbers of congruence groups", *Israel J. Math.* **88**:1-3 (1994), 31–72. MR Zbl

[Schinzel 2000] A. Schinzel, *Polynomials with special regard to reducibility*, Encycl. Math. Appl. **77**, Cambridge Univ. Press, 2000. MR Zbl

[Schmidt 1995] K. Schmidt, *Dynamical systems of algebraic origin*, Progr. Math. **128**, Birkhäuser, Basel, 1995. MR Zbl

[Schmidt and Verbitskiy 2009] K. Schmidt and E. Verbitskiy, "Abelian sandpiles and the harmonic model", *Comm. Math. Phys.* **292**:3 (2009), 721–759. MR Zbl

[Smyth 1981] C. J. Smyth, "On measures of polynomials in several variables", *Bull. Austral. Math. Soc.* **23**:1 (1981), 49–63. MR Zbl

[Stuhler 1976] U. Stuhler, "Eine Bemerkung zur Reduktionstheorie quadratischer Formen", *Arch. Math.* (*Basel*) **27**:6 (1976), 604–610. MR Zbl

[Zannier 2012] U. Zannier, *Some problems of unlikely intersections in arithmetic and geometry*, Ann. of Math. Stud. **181**, Princeton Univ. Press, 2012. MR Zbl

Communicated by Antoine Chambert-Loir
Received 2019-10-01 Revised 2023-05-25 Accepted 2023-11-27

vesselin.dimitrov@gmail.com *School of Mathematics, Georgia Institute of Technology, Atlanta, GA, United States*

philipp.habegger@unibas.ch *Departement Mathematik und Informatik, Universität Basel, Basel, Switzerland*

ALGEBRA AND NUMBER THEORY 18:11 (2024)
https://doi.org/10.2140/ant.2024.18.2003

Rooted tree maps for multiple L-values from a perspective of harmonic algebras

Hideki Murahara, Tatsushi Tanaka and Noriko Wakabayashi

We show the image of rooted tree maps forms a subspace of the kernel of the evaluation map of multiple L-values. To prove this, we define the diamond product as a modified harmonic product and describe its properties. We also show that τ-conjugate rooted tree maps are their antipodes.

1. Introduction

In [8] the second author found the Connes–Kreimer's Hopf algebra of rooted trees $\mathcal{H}$ acts on $\mathfrak{H} = \mathbb{Q}\langle x, y\rangle$, the noncommutative polynomial ring in two indeterminates. We refer to the elements in $\mathrm{End}(\mathfrak{H})$ assigned to rooted trees as rooted tree maps. Rooted tree maps possess the rules coming from their coproducts. In particular, to primitive elements in $\mathcal{H}$, derivations in $\mathrm{End}(\mathfrak{H})$ are assigned. Rooted tree maps give rise to a broad class of relations (including duality, derivation relation, and Ohno's relation) among multiple zeta values. It is then shown that the class of relations coming from rooted tree maps is equivalent to the linear part of Kawashima's relation [2]. A pending issue is that the map assigned to the antipode of a rooted tree is nothing but the conjugation of the original map by τ, the antiautomorphism on $\mathfrak{H}$ characterized by interchanging x and y, which is shown in [7] by using additional algebraic properties of rooted tree maps and harmonic algebras.

The second and the third authors generalize the domain of such rooted tree maps so that they must induce a broad class of relations among multiple L-values [9]. Unlike the case of multiple zeta values, they only show that the maps assigned to the antipodes of rooted trees induce relations among multiple L-values. We show in this paper that their first prospect is true owing to further algebraic properties of rooted tree maps and harmonic algebras to establish the basics of rooted tree maps for multiple L-values.

To be more precise, let μ_r be the set of r-th roots of unity. For an index set $(\boldsymbol{k}; \boldsymbol{s}) = (k_1, \dots, k_l; s_1, \dots, s_l)$ with $k_1, \dots, k_l \geq 1$, $s_1, \dots, s_l \in \mu_r$, $(k_1, s_1) \neq (1, 1)$, the multiple L-value of shuffle type (abbreviated as MLV) is defined in [1] by the convergent series

$$L(\boldsymbol{k}; \boldsymbol{s}) = \lim_{m \to \infty} \sum_{m > m_1 > \cdots > m_l > 0} \frac{s_1^{m_1 - m_2} \cdots s_{l-1}^{m_{l-1} - m_l} s_l^{m_l}}{m_1^{k_1} \cdots m_l^{k_l}}.$$

MSC2020: primary 11M32; secondary 05C05, 16T05.
Keywords: Connes–Kreimer Hopf algebra of rooted trees, rooted tree maps, harmonic products, multiple zeta values, multiple L-values.

If $r = 1$, this is nothing but the multiple zeta value (abbreviated as MZV). The MZVs and the MLVs have been well-studied in the last three decades.

The index set $(\boldsymbol{k}; \boldsymbol{s})$ is often identified with the word $z_{\boldsymbol{k},\boldsymbol{s}} := z_{k_1,s_1} \cdots z_{k_l,s_l}$, where $z_{k,s}$ stands for $x^{k-1}y_s$, in the noncommutative polynomial algebra $\mathcal{A}_r := \mathbb{Q}\langle x, y_s \mid s \in \mu_r \rangle$. Then MLVs are algebraically discussed via the $\mathbb{Q}$-linear map $\mathcal{L} : \mathcal{A}_r^0 \to \mathbb{C}$ defined by $\mathcal{L}(1) = 1$ and $\mathcal{L}(z_{\boldsymbol{k},\boldsymbol{s}}) = L(\boldsymbol{k}; \boldsymbol{s})$. ($\mathcal{A}_r^0$ is a subalgebra of $\mathcal{A}_r$ generated by admissible words, detailed in the next section.)

On the other hand, (nonplanar) rooted trees are finite and connected graphs with no cycles and a special vertex called the root. For example,

and so on. The topmost vertex of each rooted tree represents the root. The algebra generated by them has a Hopf algebra structure known as the Connes–Kreimer Hopf algebra of rooted trees, which appeared in [4] by Arne Dür. (One can even trace it back to the work by J. Butcher in the 1960s.) In [3], it is used in the study of perturbative quantum field theory and is well-studied in the last quarter century.

Rooted tree maps (abbreviated as RTMs), first defined in [8] based on the Connes–Kreimer Hopf algebra of rooted trees, induce a certain class of relations among MZVs. In other words, a part of $\ker \mathcal{L}$ comes from the RTMs if $r = 1$. Although this phenomenon is expected to be extended naturally to any positive integer r, the only result proved in [9] is for RTMs taken conjugation by a certain involution τ. We study some algebraic properties of RTMs for MLVs using the harmonic algebra as are studied in [7] in the MZVs case. We then show the aforementioned expectation is true and τ-conjugate RTM is nothing but its antipode.

2. Main results

Let $\mathcal{A}_r^1$ and $\mathcal{A}_r^0$ be subalgebras of $\mathcal{A}_r$ given by

$$A_r \supset A_r^1 = \mathbb{Q} \oplus A_{r,+}^1 \supset A_r^0 = \mathbb{Q} \oplus A_{r,+}^0,$$

where

$$A_{r,+}^1 = \bigoplus_{s \in \mu_r} \mathcal{A}_r y_s, \quad A_{r,+}^0 = \bigoplus_{s \in \mu_r} x \mathcal{A}_r y_s \oplus \bigoplus_{\substack{s,t \in \mu_r \\ t \neq 1}} y_t \mathcal{A}_r y_s.$$

Each word $z_{\boldsymbol{k},\boldsymbol{s}} \in \mathcal{A}_{r,+}^0$ is called admissible and corresponds to the index set $(\boldsymbol{k}; \boldsymbol{s})$ with $(k_1, s_1) \neq (1, 1)$. Let $z = x + y_1$, $z_s^\delta = x + \delta(s) y_s \in \mathcal{A}_r$, where $\delta(1) = 0$ and $\delta(s) = 1$ if $s \neq 1$.

Denote by $\mathcal{H}$ the $\mathbb{Q}$-vector space generated by rooted forests, i.e., disjoint unions of rooted trees. This $\mathcal{H}$ has a structure of a connected Hopf algebra, which is briefly described in the next section. We assign to any rooted tree t a linear map $\tilde{t} \in \mathrm{End}_{\mathbb{Q}}(\mathcal{A}_r)$, which we call a RTM, elaborated in Section 4. The assignment $\tilde{\ }$ is known to be an algebra homomorphism, and hence we can assign to any $f \in \mathcal{H}$ a linear map $\tilde{f} \in \mathrm{End}_{\mathbb{Q}}(\mathcal{A}_r)$. Using the notation of the diamond product $\diamond_s$ ($s \in \mu_r$), which is described in Section 5, we have the following result.

Theorem 2.1. *For $f \in \mathcal{H}$, there exists a unique $F_f \in \mathcal{A}_1^1$ such that*

$$\tilde{f}(z_s^\delta w) = z_s^\delta (F_f \diamond_s w)$$

for any $s \in \mu_r$ and any $w \in \mathcal{A}_r$.

The product $\diamond_s$ is a variation of the harmonic product. Indeed, Proposition 5.4 below asserts that

$$v \diamond_s w = \psi_s(\varphi(v) * \psi_s^{-1}(w)), \tag{1}$$

where $v \in \mathcal{A}_1$, $w \in \mathcal{A}_r$, and $*$ is the harmonic product. Here, $\psi_s = \varphi \mathcal{I} M_s$, where φ is the automorphism on $\mathcal{A}_r$ determined by $\varphi(x) = z$ and $\varphi(y_s) = z_s^\delta - z (= \delta(s) y_s - y_1)$ for $s \in \mu_r$, and $\mathcal{I}$ and $M_s (s \in \mu_r)$ are linear maps on $\mathcal{A}_r$ defined by

$$\mathcal{I}(\boldsymbol{z}_{\boldsymbol{k},\boldsymbol{s}} x^a) = z_{k_1,s_1} z_{k_2,s_1 s_2} \cdots z_{k_l,s_1 \cdots s_l} x^a, \quad M_s(\boldsymbol{z}_{\boldsymbol{k},\boldsymbol{s}} x^a) = z_{k_1,s s_1} z_{k_2,s_2} \cdots z_{k_l,s_l} x^a$$

for $a \geq 0$. Note that φ is an involution. According to [6], we have

$$z_s^\delta \cdot \psi_s(\mathcal{A}_{1,+}^1 * \mathcal{A}_{r,+}^1) \subset \ker \mathcal{L} \tag{2}$$

for any $s \in \mu_r$. Hence, for $s \in \mu_r$, $w \in \mathcal{A}_{r,+}^1$, and $f \in \mathrm{Aug}(\mathcal{H})$, where $\mathrm{Aug}(\mathcal{H})$ denotes the augmentation ideal of $\mathcal{H}$, i.e., $\mathcal{H} = \mathbb{Q} \oplus \mathrm{Aug}(\mathcal{H})$, we have

$$\tilde{f}(z_s^\delta w) = z_s^\delta (F_f \diamond_s w) = z_s^\delta \cdot \psi_s(\varphi(F_f) * \psi_s^{-1}(w)) \in \ker \mathcal{L}.$$

Thus we have the following:

Corollary 2.2. *For $f \in \mathrm{Aug}(\mathcal{H})$, we have $\tilde{f}(\mathcal{A}_{r,+}^0) \subset \ker \mathcal{L}$.*

Remark 2.3. This result was expected but not proved in [9]. Still we do not know the way to prove this directly from the definition of RTM (except that the case of $r = 1$, the MZV case, which is done in [8]).

Let S be the antipode of $\mathcal{H}$. Then, for $f \in \mathcal{H}$, we find that the antipode $\widetilde{S(f)}$ is described similarly by using the diamond product $\diamond_s$.

Theorem 2.4. *For $f \in \mathcal{H}$, there exists a unique $G_f \in \mathcal{A}_1^1$ such that*

$$\widetilde{S(f)}(z_s^\delta w) = z_s^\delta (G_f \diamond_s w)$$

for any $s \in \mu_r$ and any $w \in \mathcal{A}_r$.

As is defined in [9], let τ be the antiautomorphism on $\mathcal{A}_r$ defined by $\tau(x) = y_1$, $\tau(y_1) = x$, and $\tau(y_s) = -y_s$ $(s \neq 1)$. Note that τ is an involution. Then we show the following result, which is a generalization of [7, Theorem 1.5].

Theorem 2.5. *For $f \in \mathcal{H}$, we have $\widetilde{S(f)} = \tau \tilde{f} \tau$.*

Hence, for $s \in \mu_r$, $w \in \mathcal{A}^1_{r,+}$, and $f \in \mathrm{Aug}(\mathcal{H})$, we have

$$\tau \tilde{f} \tau(z_s^\delta w) = \widetilde{S(f)}(z_s^\delta w) = z_s^\delta (G_f \diamond_s w) = z_s^\delta \cdot \psi_s(\varphi(G_f) * \psi_s^{-1}(w)) \in \ker \mathcal{L}$$

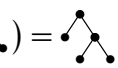
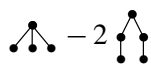
because of (1), (2), and Theorems 2.4 and 2.5. Thus again we have the following, proved first in [9, Theorem 2.4]:

Corollary 2.6. *For $f \in \mathrm{Aug}(\mathcal{H})$, we have $\tau \tilde{f} \tau(\mathcal{A}^0_{r,+}) \subset \ker \mathcal{L}$.*

3. Connes–Kreimer Hopf algebra of rooted trees

We briefly review the Connes–Kreimer Hopf algebra of rooted trees [3]. A rooted tree is a finite, connected, acyclic, and oriented graph with a special vertex called the root from which every edge directly or indirectly originates. A rooted forest is a product (disjoint union) of rooted trees. The empty forest (with no tree in it) denoted by $\mathbb{I}$ is the neutral element for the product. We denote by $\mathcal{T}$ the $\mathbb{Q}$-vector space freely generated by rooted trees.

As is mentioned in the previous section, we denote by $\mathcal{H}$ the $\mathbb{Q}$-algebra generated by rooted trees. As a vector space, $\mathcal{H}$ is freely generated by rooted forests. The $\mathbb{Q}$-linear map called the grafting operator $B_+ : \mathcal{H} \to \mathcal{T}$ is defined by $B_+(\mathbb{I}) = \bullet$ and, for a rooted forest f of positive degree, all the roots of connected components of f are grafted to a single new vertex, which becomes the new root. For example, we have

$B_+\big(\bullet\ \ \big) =$  , $\quad B_+\big(\bullet\bullet\bullet - 2\ \ \big) =$  .

In particular, the map B_+ increases the degree of the graph by 1.

We define the coproduct Δ on $\mathcal{H}$ recursively by multiplicativity and

$$\Delta(t) = t \otimes \mathbb{I} + (\mathbb{I} \otimes B_+)\Delta(f) \tag{3}$$

for $t = B_+(f)$. In terms of Hochschild cohomology of bialgebras, the grafting operator B_+ satisfies the Hochschild 1-cocycle condition. For example, we have

$$\begin{aligned}
\Delta(\mathbb{I}) &= \mathbb{I} \otimes \mathbb{I}, \\
\Delta(\bullet) &= \bullet \otimes \mathbb{I} + \mathbb{I} \otimes \bullet, \\
\Delta(\bullet\bullet) &= \bullet\bullet \otimes \mathbb{I} + 2\,\bullet \otimes \bullet + \mathbb{I} \otimes \bullet\bullet, \\
\Delta\big(\ \ \big) &= \ \ \otimes \mathbb{I} + \bullet \otimes \bullet + \mathbb{I} \otimes \ \ , \\
\Delta\big(\ \ \big) &= \ \ \otimes \mathbb{I} + \bullet\bullet \otimes \bullet + 2\,\bullet \otimes \ \ + \mathbb{I} \otimes \ \ .
\end{aligned}$$

It is known that the coproduct Δ is coassociative but not cocommutative.

The counit $\hat{\mathbb{I}} : \mathcal{H} \to \mathbb{Q}$ is defined by vanishing on $\mathrm{Aug}(\mathcal{H})$ and $\hat{\mathbb{I}}(\mathbb{I}) = 1$. If we denote the product by $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$, we define the antipode S by the antiautomorphism on $\mathcal{H}$ satisfying

$$m \circ (S \otimes \mathrm{id}) \circ \Delta = \mathbb{I} \circ \hat{\mathbb{I}} = m \circ (\mathrm{id} \otimes S) \circ \Delta .$$

Then the tuple $(\mathcal{H}, m, \mathbb{I}, \Delta, \hat{\mathbb{I}}, S)$ forms a Hopf algebra known as the Connes–Kreimer Hopf algebra of rooted trees.

4. Rooted tree maps

We introduce rooted tree maps developed in [9]. Let the identity map on $\mathcal{A}_r$ be assigned to the empty forest $\mathbb{I}$, i.e., $\tilde{\mathbb{I}} = \mathrm{id}$. For any rooted forest f of positive degree, we define the $\mathbb{Q}$-linear map $\tilde{f} : \mathcal{A}_r \to \mathcal{A}_r$ by the following four conditions:

(I) if $f = \bullet$, $\tilde{f}(z_s^\delta) = z_s^\delta(z - z_s^\delta)$ and $\tilde{f}(z) = 0$,

(II) $\widetilde{B_+(f)}(z_s^\delta) = R_{z-z_s^\delta} R_{2z-z_s^\delta} R_{z-z_s^\delta}^{-1} \tilde{f}(z_s^\delta)$ and $\widetilde{B_+(f)}(z) = 0$,

(III) if $f = gh$, $\tilde{f}(v) = \tilde{g}(\tilde{h}(v))$ for $v \in \{z, z_s^\delta \mid s \in \mu_r\}$,

(IV) $\tilde{f}(wv) = M(\widetilde{\Delta(f)}(w \otimes v))$ for $w \in \mathcal{A}_r$, $v \in \{z, z_s^\delta \mid s \in \mu_r\}$,

where $s \in \mu_r$, R_w denotes the right multiplication map by w, that is, $R_w(v) = vw$ for $v, w \in \mathcal{A}_r$, $M : \mathcal{A}_r \otimes \mathcal{A}_r \to \mathcal{A}_r$ denotes the concatenation product, and $\widetilde{\Delta(f)} = \sum_{(f)} \widetilde{f'} \otimes \widetilde{f''}$ when $\Delta(f) = \sum_{(f)} f' \otimes f''$. As a matter of fact, the assignment $\tilde{\cdot} : \mathcal{H} \to \mathrm{End}_{\mathbb{Q}}(\mathcal{A}_r)$ is an algebra homomorphism. We find that $\tilde{f}(z_s^\delta)$ always ends with $z - z_s^\delta$, and hence the condition (II) is well-defined. We also find that the image $\tilde{f}(v)$ in the condition (III) does not depend on how to decompose f into g and h because of the commutativity property of RTMs which is proved by induction on graph order. We can also show the conditions (III) and (IV) hold for any $v \in \mathcal{A}_r$; see [8, Theorem 1.2] or [9, Theorem 2.2]. We call $\tilde{f}$ the RTM assigned to $f \in \mathcal{H}$.

Example 4.1 (calculations of images of RTMs). Since $\tilde{\bullet}(z_s^\delta) = z_s^\delta(z - z_s^\delta)$ and $\Delta(\bullet) = \bullet \otimes \mathbb{I} + \mathbb{I} \otimes \bullet$,

$$\tilde{\bullet}\tilde{\bullet}(z_s^\delta) = \tilde{\bullet}(z_s^\delta(z - z_s^\delta)) = \tilde{\bullet}(z_s^\delta)(z - z_s^\delta) + z_s^\delta \tilde{\bullet}(z - z_s^\delta) = z_s^\delta(z - z_s^\delta)^2 + (z_s^\delta)^2(z - z_s^\delta).$$

Then we calculate

$$\curlywedge(z_s^\delta) = \widetilde{B_+(\bullet\bullet)}(z_s^\delta) = R_{z-z_s^\delta} R_{2z-z_s^\delta} R_{z-z_s^\delta}^{-1} \tilde{\bullet}\tilde{\bullet}(z_s^\delta) = z_s^\delta(z-z_s^\delta)(2z-z_s^\delta)(z-z_s^\delta) + (z_s^\delta)^2(2z-z_s^\delta)(z-z_s^\delta).$$

5. Harmonic product and diamond product

The harmonic product $* : \mathcal{A}_r \times \mathcal{A}_r \to \mathcal{A}_r$ is defined by $\mathbb{Q}$-bilinearity and

(I) $1 * w = w * 1 = w$,

(II) $vy_s * wy_t = (v * wy_t)y_s + (vy_s * w)y_t + (v * w)xy_{st}$,

(III) $vx * w = v * wx = (v * w)x$

for $v, w \in \mathcal{A}_r$, $s, t \in \mu_r$. It is associative and commutative. The tuples $(\mathcal{A}_r^1, *)$ and $(\mathcal{A}_r^0, *)$ are subalgebras of $(\mathcal{A}_r, *)$. The composition $\mathcal{LI}$ is known as the evaluation map of MLVs of harmonic type. It is an algebra homomorphism with respect to $*$ (see [1]).

Lemma 5.1. *For $k, l \geq 1$, $s, t \in \mu_r$, and $v, w \in \mathcal{A}_r$, we have*

(i) $vz_{k,s} * wz_{l,t} = (v * wz_{l,t})z_{k,s} + (vz_{k,s} * w)z_{l,t} + (v * w)z_{k+l,st}$,

(ii) $z_{k,s}v * z_{l,t}w = z_{k,s}(v * z_{l,t}w) + z_{l,t}(z_{k,s}v * w) + z_{k+l,st}(v * w)$.

Proof. Because of the condition (III), it is enough to show when $v, w \in \mathcal{A}_r^1$.

To show (i), substitute vx^{k-1} and wx^{l-1} into v and w, respectively, in the condition (II) and then use the condition (III).

We show (ii) by induction on total degree of words. If $v = w = 1$, it follows from (i) for $v = w = 1$.

If $v = v'z_{m,a}$ ($v' \in \mathcal{A}_r^1$, $m \geq 1$, $a \in \mu_r$) and $w = 1$, the left-hand side equals

$$(z_{k,s}v' * z_{l,t})z_{m,a} + z_{k,s}vz_{l,t} + z_{k,s}v'z_{l+m,ta} \tag{4}$$

because of (i). The first term turns into

$$(z_{k,s}(v' * z_{l,t}) + z_{l,t}z_{k,s}v' + z_{k+l,st}v')z_{m,a}$$

by induction, and hence we have

$$(4) = z_{k,s}((v' * z_{l,t})z_{m,a} + vz_{l,t} + v'z_{l+m,ta}) + z_{l,t}z_{k,s}v + z_{k+l,st}v.$$

Again by (i), we see that this coincides with the right-hand side. The proof goes similarly if $v = 1$ and $w = w'z_{n,b}$ ($w' \in \mathcal{A}_r$, $n \geq 1$, $b \in \mu_r$).

If $v = v'z_{m,a}$ and $w = w'z_{n,b}$, the left-hand side equals

$$(z_{k,s}v' * z_{l,t}w)z_{m,a} + (z_{k,s}v * z_{l,t}w')z_{n,b} + (z_{k,s}v' * z_{l,t}w')z_{m+n,ab}$$

because of (i). This turns into

$$\begin{aligned}(z_{k,s}(v' * z_{l,t}w) + z_{l,t}(z_{k,s}v' * w) + z_{k+l,st}(v' * w))z_{m,a}\\ + (z_{k,s}(v * z_{l,t}w') + z_{l,t}(z_{k,s}v * w') + z_{k+l,st}(v * w'))z_{n,b}\\ + (z_{k,s}(v' * z_{l,t}w') + z_{l,t}(z_{k,s}v' * w') + z_{k+l,st}(v' * w'))z_{m+n,ab}\end{aligned}$$

by induction. Again by (i), we see that this coincides with the right-hand side. □

From now on, let $y = y_1$ for simplicity. For $s \in \mu_r$, we define the $\mathbb{Q}$-bilinear map $\diamond_s : \mathcal{A}_1 \times \mathcal{A}_r \to \mathcal{A}_r$ by

$$\begin{aligned}1 \diamond_s w &= w,\\ v \diamond_s 1 &= \psi_s\varphi(v),\\ vx \diamond_s wx &= (v \diamond_s wx)x - (vy \diamond_s w)x,\\ vy \diamond_s wx &= (v \diamond_s wx)y + (vy \diamond_s w)x,\\ vx \diamond_s wy &= (v \diamond_s wy)x + (vx \diamond_s w)y,\\ vy \diamond_s wy &= (v \diamond_s wy)y - (vx \diamond_s w)y,\\ vx \diamond_s wy_t &= (v \diamond_s wy_t)x + (v \diamond_s wz_t)y_t - (vy \diamond_s w)y_t,\\ vy \diamond_s wy_t &= (v \diamond_s wy_t)y - (v \diamond_s wz_t)y_t + (vy \diamond_s w)y_t,\end{aligned} \tag{5}$$

for $v \in \mathcal{A}_1$, $w \in \mathcal{A}_r$ and $1 \neq t \in \mu_r$. When $r = 1$, the product $\diamond_1$ corresponds to the one defined in [5] and is commutative. In general, 1 is the left unit but not the right unit. For example, one checks $y \diamond_s 1 = z - z_s^\delta$.

Lemma 5.2. *For $s \in \mu_r$, $v \in \mathcal{A}_1$, and $w \in \mathcal{A}_r$, we have*

$$vz \diamond_s w = v \diamond_s wz = (v \diamond_s w)z.$$

Proof. By definition, we easily see $vz \diamond_s w = (v \diamond_s w)z$.

We prove $v \diamond_s wz = (v \diamond_s w)z$ for words v, w by induction on $d = \deg(v)$. It is obvious if $d = 0$. Assume $d \geq 1$. If $v = v'x$, by definition (in particular, adding the third and fifth identities in (5)), we have

$$\begin{aligned} v'x \diamond_s wz &= (v' \diamond_s wz)x - (v'y \diamond_s w)x + (v'x \diamond_s w)y \\ &= (v' \diamond_s wz)x - (v'z \diamond_s w)x + (v'x \diamond_s w)z. \end{aligned}$$

By the induction hypothesis, the first two terms cancel out, and hence we obtain the assertion. The proof goes similarly when $v = v'y$. □

Lemma 5.3. *For $s, t \in \mu_r$, $v \in \mathcal{A}_1$, and $w \in \mathcal{A}_r$, we have*

(i) $vx \diamond_s wz_t^\delta = (v \diamond_s wz_t^\delta)z_t^\delta - (vy \diamond_s w)z_t^\delta$,

(ii) $vy \diamond_s wz_t^\delta = (v \diamond_s wz_t^\delta)(y \diamond_t 1) + (vy \diamond_s w)z_t^\delta$.

Proof. By the third and seventh identities in (5), we have (i). By (i), Lemma 5.2, and $y \diamond_t 1 = z - z_t^\delta$, we have (ii). □

We put $z_s = x + y_s$ for simplicity (and hence $z_1 = z$). Note that $\varphi(z_s) = z_s^\delta$.

Proposition 5.4. *For $s \in \mu_r$, $v \in \mathcal{A}_1$, and $w \in \mathcal{A}_r$, we have*

$$v \diamond_s w = \psi_s\big(\varphi(v) * \psi_s^{-1}(w)\big).$$

Proof. If $v = 1$ or $w = 1$, it is obvious. Otherwise, the proof goes by induction on $\deg(v) + \deg(w)$.

If $v = v'z$, by definitions, the right-hand side turns into

$$\psi_s(\varphi(v'z) * \psi_s^{-1}(w)) = \psi_s(\varphi(v')x * \psi_s^{-1}(w)) = \psi_s\big((\varphi(v') * \psi_s^{-1}(w))x\big) = \psi_s(\varphi(v') * \psi_s^{-1}(w))z.$$

Then, by the induction hypothesis, this equals $(v' \diamond_s w)z$, which equals the left-hand side because of Lemma 5.2.

Similarly, if $w = w'z$, the right-hand side turns into

$$\psi_s(\varphi(v) * \psi_s^{-1}(w'z)) = \psi_s(\varphi(v) * \psi_s^{-1}(w')x) = \psi_s\big((\varphi(v) * \psi_s^{-1}(w'))x\big) = \psi_s(\varphi(v) * \psi_s^{-1}(w'))z,$$

which equals the left-hand side.

To complete the proof, we show when $v = v'x$ and $w = w'z_t^\delta$. In this case, the right-hand side turns into

$$\psi_s(\varphi(v')z * \psi_s^{-1}(w'z_t^\delta)). \tag{6}$$

Without loss of generality, suppose $\varphi(w') = z_{k_1,t_1} \cdots z_{k_n,t_n}$. Then, by definitions, we find

$$\psi_s^{-1}(w' z_t^\delta) = \psi_s^{-1}(w') z_{t/t_n}.$$

Hence, we have

$$(6) = \psi_s\big((\varphi(v')y * \psi_s^{-1}(w'))z_{t/t_n} + (\varphi(v') * \psi_s^{-1}(w')y_{t/t_n})z + (\varphi(v') * \psi_s^{-1}(w'))x z_{t/t_n}\big). \tag{7}$$

Since $\varphi(v') \in \mathcal{A}_1$, $\psi_s^{-1}(w') = z_{k_1,t_1/s} z_{k_2,t_2/t_1} \cdots z_{k_n,t_n/t_{n-1}}$, and the harmonic product has combinatorial meaning of overlapping shuffle, the subscript of 'y' in the last z_{t/t_n} or z changes into

$$s \times \left(\frac{t_1}{s} \times \frac{t_2}{t_1} \times \cdots \times \frac{t_n}{t_{n-1}}\right) \times \frac{t}{t_n} = t \quad \text{or} \quad s \times \left(\frac{t_1}{s} \times \frac{t_2}{t_1} \times \cdots \times \frac{t_n}{t_{n-1}} \times \frac{t}{t_n}\right) = t,$$

respectively, after the map ψ_s applies. Therefore we have

$$\begin{aligned}(7) &= \psi_s\big(\varphi(v')y * \psi_s^{-1}(w') + \varphi(v') * \psi_s^{-1}(w')y_{t/t_n} + (\varphi(v') * \psi_s^{-1}(w'))x\big)z_t^\delta \\ &= \psi_s(-\varphi(v'y) * \psi_s^{-1}(w') + \varphi(v') * \psi_s^{-1}(w' z_t^\delta))z_t^\delta \\ &= (-v'y \diamond_s w' + v' \diamond_s w' z_t^\delta) z_t^\delta.\end{aligned}$$

The last equality is by the induction hypothesis. By Lemma 5.3(i), this coincides with the left-hand side. □

Lemma 5.5. *The product $\diamond_s$ gives a left $\mathcal{A}_1$-module structure to $\mathcal{A}_r$ for any $s \in \mu_r$.*

Proof. For $s \in \mu_r$, $u, v \in \mathcal{A}_1$ and $w \in \mathcal{A}_r$, We have

$$\begin{aligned}(u \diamond_1 v) \diamond_s w &= \big(\varphi(\varphi(u) * \varphi(v))\big) \diamond_s w \\ &= \psi_s\big((\varphi(u) * \varphi(v)) * \psi_s^{-1}(w)\big) \\ &= \psi_s\big(\varphi(u) * (\varphi(v) * \psi_s^{-1}(w))\big) \\ &= \psi_s\big(\varphi(u) * (\psi_s^{-1}(v \diamond_s w))\big) = u \diamond_s (v \diamond_s w)\end{aligned}$$

by Proposition 5.4 and the associativity of $*$. □

Lemma 5.6. *For $s, t \in \mu_r$ and $v, w \in \mathcal{A}_r$, we have*

$$y \diamond_s v z_t^\delta w = (y \diamond_s v) z_t^\delta w + v z_t^\delta (y \diamond_s w) + v z_t^\delta (z_s^\delta - z_t^\delta) w.$$

Proof. We prove the lemma by induction on $\deg(w)$. When $w = 1$, we have

$$y \diamond_s v z_t^\delta = (y \diamond_s v) z_t^\delta + (1 \diamond_s v z_t^\delta)(y \diamond_t 1) \tag{8}$$

by Lemma 5.3(ii). Since $y \diamond_t 1 = y \diamond_s 1 + (z_s^\delta - z_t^\delta)$, we have

$$(8) = (y \diamond_s v) z_t^\delta + v z_t^\delta (y \diamond_s 1) + v z_t^\delta (z_s^\delta - z_t^\delta)$$

and the assertion. If $w = w'z$ ($w' \in \mathcal{A}_r$), by the induction hypothesis and Lemma 5.2, we have

$$\begin{aligned}\text{L.H.S.} &= (y \diamond_s v z_t^\delta w')z \\ &= (y \diamond_s v) z_t^\delta w' z + v z_t^\delta (y \diamond_s w') z + v z_t^\delta (z_s^\delta - z_t^\delta) w' z = \text{R.H.S.}\end{aligned}$$

If $w = w' z_{t'}^{\delta}$ ($w' \in \mathcal{A}_r$), by Lemma 5.3(ii) and the induction hypothesis, we have

$$\begin{aligned}\text{L.H.S.} &= (1 \diamond_s v z_t^{\delta} w' z_{t'}^{\delta})(y \diamond_{t'} 1) + (y \diamond_s v z_t^{\delta} w') z_{t'}^{\delta} \\ &= (1 \diamond_s v z_t^{\delta} w' z_{t'}^{\delta})(y \diamond_{t'} 1) + (y \diamond_s v) z_t^{\delta} w + v z_t^{\delta} (y \diamond_s w') z_{t'}^{\delta} + v z_t^{\delta} (z_s^{\delta} - z_t^{\delta}) w \\ &= \text{R.H.S.}\end{aligned}$$

This finishes the proof. □

Now write $R = R_y R_{x+2y} R_y^{-1}$. For rooted forests f, we define polynomials $F_f \in \mathcal{A}_1^1$ recursively by

- $F_{\mathbb{I}} = 1$,
- $F_{\bullet} = y$,
- $F_t = R(F_f)$ if $t = B_+(f)$ and $f \neq \mathbb{I}$,
- $F_f = F_g \diamond_1 F_h$ if $f = gh$.

The subscript of F is extended linearly.

Proposition 5.7. *For $f \in \mathcal{H}$, put $\Delta(f) = \sum_{(f)} f' \otimes f''$. Then, for $s, s' \in \mu_r$ and $v, w \in \mathcal{A}_r$, we have*

$$F_f \diamond_s v z_{s'}^{\delta} w = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w).$$

Proof. It is enough to consider the case that f is a monomial, i.e., a rooted forest. If $f = \mathbb{I}$, it is obvious. If $f = \bullet$, by Lemma 5.6, we find the proposition holds.

Assume $\deg(f) \geq 2$ and the proposition holds for any elements in $\mathcal{H}$ of degree less than $\deg(f)$. If $f = gh$ ($g, h \neq \mathbb{I}$), we have

$$F_f \diamond_s v z_{s'}^{\delta} w = (F_g \diamond_1 F_h) \diamond_s v z_{s'}^{\delta} w = F_g \diamond_s (F_h \diamond_s v z_{s'}^{\delta} w) \tag{9}$$

because of Lemma 5.5. Since $\deg(g), \deg(h) < \deg(f)$, we have

$$(9) = \sum_{(h)} F_g \diamond_s (F_{h'} \diamond_s v) z_{s'}^{\delta} (F_{h''} \diamond_{s'} w) = \sum_{(g)} \sum_{(h)} (F_{g'} \diamond_s (F_{h'} \diamond_s v)) z_{s'}^{\delta} (F_{g''} \diamond_{s'} (F_{h''} \diamond_{s'} w)) \tag{10}$$

by the induction hypothesis. Again by Lemma 5.5, we have

$$(10) = \sum_{(g)} \sum_{(h)} ((F_{g'} \diamond_1 F_{h'}) \diamond_s v) z_{s'}^{\delta} ((F_{g''} \diamond_1 F_{h''}) \diamond_{s'} w) = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w),$$

and hence the assertion.

If f is a tree and $f = B_+(g)$, we have $F_f = R(F_g)$. In this case, the proof goes inductively on $\deg(w)$. When $w = 1$, we have

$$\begin{aligned} F_f \diamond_s v z_{s'}^{\delta} &= R(F_g) \diamond_s v z_{s'}^{\delta} \\ &= (R_y^{-1}(F_g) x + 2F_g) y \diamond_s v z_{s'}^{\delta} \\ &= (F_f \diamond_s v) z_{s'}^{\delta} + ((R_y^{-1}(F_g) x + 2F_g) \diamond_s v z_{s'}^{\delta})(z - z_{s'}^{\delta}) \end{aligned} \tag{11}$$

because of Lemma 5.3. Since $\deg(g) < \deg(f)$, we have

$$F_g \diamond_s v z_{s'}^{\delta} = \sum_{(g)} (F_{g'} \diamond_s v) z_{s'}^{\delta} (F_{g''} \diamond_{s'} 1)$$

by the induction hypothesis. Then we have

$$\begin{aligned}(11) &= (F_f \diamond_s v) z_{s'}^{\delta} + (R_y^{-1}(F_g) x \diamond_s v z_{s'}^{\delta})(z - z_{s'}^{\delta}) + 2 \sum_{(g)} (F_{g'} \diamond_s v) z_{s'}^{\delta} (F_{g''} \diamond_{s'} 1)(z - z_{s'}^{\delta}) \\ &= (F_f \diamond_s v) z_{s'}^{\delta} + ((R_y^{-1}(F_g) \diamond_s v z_{s'}^{\delta}) z_{s'}^{\delta} - (F_g \diamond_s v) z_{s'}^{\delta})(z - z_{s'}^{\delta}) \\ &\quad + 2(F_g \diamond_s v) z_{s'}^{\delta} (y \diamond_{s'} 1) + \sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} ((R(F_{g''}) - R_y^{-1}(F_{g''}) xy) \diamond_{s'} 1) \qquad (12)\end{aligned}$$

because of Lemma 5.3(i), $y \diamond_{s'} 1 = z - z_{s'}^{\delta}$, and

$$(F_{g''} \diamond_{s'} 1)(z - z_{s'}^{\delta}) = F_{g''} y \diamond_{s'} 1 = \tfrac{1}{2}(R(F_{g''}) - R_y^{-1}(F_{g''}) xy) \diamond_{s'} 1.$$

We find

$$\begin{aligned}\sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} 1) &= \sum_{\substack{c(f) \\ f'' \neq \mathbb{1}}} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} 1) + (F_f \diamond_s v) z_{s'}^{\delta} \\ &= \sum_{(g)} (F_{g'} \diamond_s v) z_{s'}^{\delta} (F_{B_+(g'')} \diamond_{s'} 1) + (F_f \diamond_s v) z_{s'}^{\delta} \\ &= \sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} (R(F_{g''}) \diamond_{s'} 1) + (F_g \diamond_s v) z_{s'}^{\delta} (y \diamond_{s'} 1) + (F_f \diamond_s v) z_{s'}^{\delta}\end{aligned}$$

because of (3). Therefore we have

$$(12) = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} 1) + (R_y^{-1}(F_g) \diamond_s v z_{s'}^{\delta}) z_{s'}^{\delta} (z - z_{s'}^{\delta}) - \sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} (R_y^{-1}(F_{g''}) xy \diamond_{s'} 1). \qquad (13)$$

We now see that the second and third terms in (13) cancel out. To see this, we need to show

$$\sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} (R_y^{-1}(F_{g''}) \diamond_{s'} 1) = R_y^{-1}(F_g) \diamond_s v z_{s'}^{\delta}$$

because of $R_y^{-1}(F_{g''}) xy \diamond_{s'} 1 = (R_y^{-1}(F_{g''}) \diamond_{s'} 1) z_{s'}^{\delta} (z - z_{s'}^{\delta})$. By $R_y^{-1}(F_{g''}) \diamond_{s'} 1 = R_{z - z_{s'}^{\delta}}^{-1}(F_{g''} \diamond_{s'} 1)$, the induction hypothesis, and Lemma 5.3, we have

$$\begin{aligned}\sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} (R_y^{-1}(F_{g''}) \diamond_{s'} 1) &= R_{z - z_{s'}^{\delta}}^{-1} \Bigg(\sum_{\substack{(g) \\ g'' \neq \mathbb{1}}} (F_{g'} \diamond_s v) z_{s'}^{\delta} (F_{g''} \diamond_{s'} 1) \Bigg) = R_{z - z_{s'}^{\delta}}^{-1} (F_g \diamond_s v z_{s'}^{\delta} - (F_g \diamond_{s'} v) z_{s'}^{\delta}) \\ &= R_{z - z_{s'}^{\delta}}^{-1} (R_y^{-1}(F_g) \diamond_s v z_{s'}^{\delta})(z - z_{s'}^{\delta}) = R_y^{-1}(F_g) \diamond_s v z_{s'}^{\delta}.\end{aligned}$$

Thus we conclude

$$(13) = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} 1).$$

Now we proceed to the case when $\deg(w) \geq 1$. If $w = w'z$ ($w' \in \mathcal{A}_r$), we have

$$F_f \diamond_s v z_{s'}^{\delta} w = (F_f \diamond_s v z_{s'}^{\delta} w') z = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w') z = \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w)$$

by Lemma 5.2 and the induction hypothesis. If $w = w' z_{s''}^{\delta}$ ($w' \in \mathcal{A}_r$, $s'' \in \mu_r$), since we have already proved the identity when $w = 1$, we have

$$\begin{aligned} F_f \diamond_s v z_{s'}^{\delta} w &= \sum_{(f)} (F_{f'} \diamond_s v z_{s'}^{\delta} w') z_{s''}^{\delta} (F_{f''} \diamond_{s''} 1) \\ &= \sum_{(f)} \sum_{(f')} (F_{f'_a} \diamond_s v) z_{s'}^{\delta} (F_{f'_b} \diamond_{s'} w) z_{s''}^{\delta} (F_{f''} \diamond_{s''} 1), \end{aligned}$$

where we put $\Delta(f') = \sum_{(f')} f'_a \otimes f'_b$. We also have

$$\begin{aligned} \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w) &= \sum_{(f)} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''} \diamond_{s'} w' z_{s''}^{\delta}) \\ &= \sum_{(f)} \sum_{(f'')} (F_{f'} \diamond_s v) z_{s'}^{\delta} (F_{f''_a} \diamond_{s'} w') z_{s''}^{\delta} (F_{f''_b} \diamond_{s''} 1), \end{aligned}$$

where we put $\Delta(f'') = \sum_{(f'')} f''_a \otimes f''_b$. By coassociativity of Δ, these two coincide, and hence we have conclusion. □

The following property plays an important role in our proof of Theorem 2.5 in Section 8.

Proposition 5.8. *For $s, t \in \mu_r$, $v \in \mathcal{A}_1$, and $w \in \mathcal{A}_r$, we have*

$$z_s^{\delta}(vy \diamond_s w(z - z_t^{\delta})) = -\tau(\tau(v) y \diamond_t \tau(w)(z - z_s^{\delta}))(z - z_t^{\delta}).$$

Proof. By Lemmas 5.3(ii) and 5.9(ii), it is equivalent to show the identity

$$vy \diamond_s w - v \diamond_s w z_t^{\delta} = \tau(\tau(v) \diamond_t \tau(w) z_s^{\delta} - \tau(v) y \diamond_t \tau(w)). \tag{14}$$

We prove this by induction on $\deg(v) + \deg(w)$. We consider five cases.

Case 1: $v = 1$. If $w = 1, z, z - z_u^{\delta}$ ($u \in \mu_r$), we calculate that both sides turn into

$$z - z_s^{\delta} - z_t^{\delta}, \quad z(z - z_t^{\delta}) - z_s^{\delta} z, \quad (z - z_s^{\delta} - z_u^{\delta})(z - z_u^{\delta}) - (z - z_u^{\delta}) z_t^{\delta},$$

respectively. If $w = w'z$, we calculate

$$\begin{aligned} \tau(\text{R.H.S.}) &= z\tau(w') z_s^{\delta} - (y \diamond_t 1)(1 \diamond_t z\tau(w')) - z(y \diamond_t \tau(w')) + z(y \diamond_t 1)(1 \diamond_t \tau(w')) \\ &= \tau((y \diamond_s w') z) - \tau(w' z z_t^{\delta}) = \tau(\text{L.H.S.}) \end{aligned}$$

by Lemma 5.9(iii) and the induction hypothesis. If $w = w'(z - z_u^\delta)$, we have

$$\text{L.H.S.} = (y \diamond_s w')(z - z_u^\delta) - w' z_u^\delta (z - z_u^\delta) - w'(z - z_u^\delta) z_t^\delta$$

by Lemma 5.9(ii), and

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_u^\delta \tau(w') z_s^\delta - (y \diamond_t 1) z_u^\delta \tau(w') - z_u^\delta (y \diamond_u \tau(w')) \\ &= z_u^\delta \tau(y \diamond_s w' - 1 \diamond_s w' z_u^\delta) - (y \diamond_t 1) z_u^\delta \tau(w')\end{aligned}$$

by Lemma 5.9(iv) and the induction hypothesis. Thus (14) holds.

Case 2: $v = z$. If $w = 1, z, z - z_u^\delta$, we calculate that both sides turn into

$$z^2 - z z_s^\delta - z_t^\delta z, \quad z(z - z_s^\delta - z_t^\delta) z, \quad (z^2 - z z_s^\delta - z_u^\delta z)(z - z_u^\delta) - (z - z_u^\delta) z_t^\delta z,$$

respectively. If $w = w'z$, we calculate

$$\begin{aligned}\tau(\text{R.H.S.}) &= z(z \diamond_t \tau(w') z_s^\delta) - z\big(y \diamond_t \tau(w) + z y \diamond_t \tau(w') - z(y \diamond_t \tau(w'))\big) \\ &= z\tau(z y \diamond_s w' - z \diamond_s w' z_t^\delta) - z\big(y \diamond_t \tau(w) - z(y \diamond_t \tau(w'))\big)\end{aligned}$$

by Lemma 5.9(v) and the induction hypothesis. Applying Lemma 5.9(iii) to the third term, we find that this is $\tau(\text{L.H.S.})$. Note that

$$x v \diamond_s z_t^\delta w = z_s^\delta (v \diamond_s z_t^\delta w) + z_t^\delta (x v \diamond_t w) - z z_t^\delta (v \diamond_t w) \tag{15}$$

by Lemma 5.9(iv) and (vi). If $w = w'(z - z_u^\delta)$, we have

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_u^\delta (z \diamond_u \tau(w') z_s^\delta) - z(y \diamond_t z_u^\delta \tau(w')) - z_u^\delta (z y \diamond_u \tau(w')) + z z_u^\delta (y \diamond_u \tau(w')) \\ &= z_u^\delta \tau(z y \diamond_s w' - z \diamond_s w' z_u^\delta) - z(y \diamond_t 1)(1 \diamond_t z_u^\delta \tau(w'))\end{aligned}$$

by (15) and the induction hypothesis. This is $\tau(\text{L.H.S.})$ because of Lemma 5.9(ii). Thus (14) holds.

Case 3: $v = y$. If $w = 1, z, z - z_u^\delta$, we calculate that both sides turn into

$$\begin{aligned}&(z - z_s^\delta - z_t^\delta)(z - z_t^\delta) - (z - z_s^\delta) z_s^\delta, \quad (z - z_s^\delta)^2 z - (z - z_s^\delta) z z_t^\delta - z z_t^\delta (z - z_t^\delta), \\ &(z - z_s^\delta - z_u^\delta)(z - z_u^\delta)(z - z_t^\delta - z_u^\delta) - (z - z_s^\delta) z_s^\delta (z - z_u^\delta) - (z - z_u^\delta) z_t^\delta (z - z_t^\delta),\end{aligned}$$

respectively. Note that

$$x v \diamond_s z w = z_s^\delta (v \diamond_s z w) + z(x v \diamond_s w) - z z_s^\delta (v \diamond_s w) \tag{16}$$

by Lemma 5.9(iii) and (v). If $w = w'z$, we calculate

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta (1 \diamond_t z \tau(w') z_s^\delta) + z(x \diamond_t \tau(w') z_s^\delta) - z z_t^\delta (1 \diamond_t \tau(w') z_s^\delta) \\ &\qquad - \big(z_t^\delta (y \diamond_t z \tau(w')) + z(x y \diamond_t \tau(w')) - z z_t^\delta (y \diamond_t \tau(w'))\big) \\ &= z_t^\delta \tau(y \diamond_s w - 1 \diamond_s w z_t^\delta) + z\tau(y^2 \diamond_s w' - y \diamond_s w' z_t^\delta) - z z_t^\delta (y \diamond_s w' - 1 \diamond_s w' z_t^\delta)\end{aligned}$$

by (16) and the induction hypothesis. Hence, applying τ and using Lemma 5.3(ii), we find (14) also holds in this case. If $w = w'(z - z_u^\delta)$, we have

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(1 \diamond_t \tau(w) z_s^\delta) + z_u^\delta(x \diamond_u \tau(w') z_s^\delta) - z z_u^\delta(1 \diamond_u \tau(w') z_s^\delta) \\ &\quad - \big(z_t^\delta(y \diamond_t \tau(w)) + z_u^\delta(xy \diamond_u \tau(w')) - z z_u^\delta(y \diamond_u \tau(w'))\big) \\ &= z_t^\delta \tau(y \diamond_s w - 1 \diamond_s w z_t^\delta) + z_u^\delta \tau(y^2 \diamond_s w' - y \diamond_s w' z_u^\delta) - z z_u^\delta \tau(y \diamond_s w' - 1 \diamond_s w' z_u^\delta)\end{aligned}$$

by (15) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.9(ii) and 5.3(ii). Thus (14) holds.

Case 4: $v = v'z$. If $w = 1$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z(\tau(v') \diamond_t z_s^\delta) + (z_s^\delta z - z z_s^\delta)(\tau(v') \diamond_s 1) - z(\tau(v') y \diamond_t 1) \\ &= z_s^\delta z \tau(v' \diamond_s 1) - \tau(v \diamond_s z_t^\delta)\end{aligned}$$

by Lemma 5.9(i), (vi), and (vii) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.9(i). If $w = z$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z\big(\tau(v') \diamond_t z z_s^\delta + \tau(v) \diamond_t z_s^\delta - z(\tau(v') \diamond_t z_s^\delta)\big) - z\big(\tau(v') y \diamond_t z + \tau(v) y \diamond_t 1 - z(\tau(v') y \diamond_t 1)\big) \\ &= z\tau(v' y \diamond_s z - v' \diamond_s z z_t^\delta) + z\tau(vy \diamond_s 1 - v \diamond_s z_t^\delta) - z\tau(v' y \diamond_s 1 - v' \diamond_s z_t^\delta)\end{aligned}$$

by Lemma 5.9(v) and the induction hypothesis. This is τ(L.H.S.) because of Lemma 5.2. If $w = z - z_u^\delta$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z(\tau(v') \diamond_t z_u^\delta z_s^\delta) + z_u^\delta(\tau(v) \diamond_u z_s^\delta) - z z_u^\delta(\tau(v') \diamond_u z_s^\delta) \\ &\quad - \big(z(\tau(v') y \diamond_t z_u^\delta) + z_u^\delta(\tau(v) y \diamond_u 1) - z z_u^\delta(\tau(v') y \diamond_u 1)\big) \\ &= z\tau(v' y \diamond_s (z - z_u^\delta) - v' \diamond_s (z - z_u^\delta) z_t^\delta) + z_u^\delta \tau(vy \diamond_s 1 - v \diamond_s z_u^\delta) - z z_u^\delta \tau(v' y \diamond_s 1 - v' \diamond_s z_u^\delta)\end{aligned}$$

by Lemma 5.9(vi) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). If $w = w'z$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z\big(\tau(v') \diamond_t \tau(w) z_s^\delta + \tau(v) \diamond_t \tau(w') z_s^\delta - z(\tau(v') \diamond_t \tau(w') z_s^\delta)\big) \\ &\quad - z\big(\tau(v') y \diamond_t \tau(w) + \tau(v) y \diamond_t \tau(w') - z(\tau(v') y \diamond_t \tau(w'))\big) \\ &= z\tau(v' y \diamond_s w - v' \diamond_s w z_t^\delta) + z\tau(vy \diamond_s w' - v \diamond_s w' z_t^\delta) - z\tau(v' y \diamond_s w' - v' \diamond_s w' z_t^\delta)\end{aligned}$$

by Lemma 5.9(v) and the induction hypothesis. This is τ(L.H.S.) because of Lemma 5.2. If $w = w'(z - z_u^\delta)$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z(\tau(v') \diamond_t \tau(w) z_s^\delta) + z_u^\delta(\tau(v) \diamond_u \tau(w') z_s^\delta) - z z_u^\delta(\tau(v') \diamond_u \tau(w') z_s^\delta) \\ &\quad - \big(z(\tau(v') y \diamond_t \tau(w)) + z_u^\delta(\tau(v) y \diamond_u \tau(w')) - z z_u^\delta(\tau(v') y \diamond_u \tau(w'))\big) \\ &= z\tau(v' y \diamond_s w - v' \diamond_s w z_t^\delta) + z_u^\delta \tau(vy \diamond_s w' - v \diamond_s w' z_u^\delta) - z z_u^\delta \tau(v' y \diamond_s w' - v' \diamond_s w' z_u^\delta)\end{aligned}$$

by Lemma 5.9(vi) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). Thus (14) holds.

Case 5: $v = v'y$. If $w = 1$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(\tau(v') \diamond_t z_s^\delta) + z_s^\delta(\tau(v) \diamond_s 1) - zz_s^\delta(\tau(v') \diamond_s 1) - z_t^\delta(\tau(v')y \diamond_t 1) \\ &= \tau(v'y \diamond_s 1 - v' \diamond_s z_t^\delta) + z_s^\delta(\tau(v) \diamond_s 1) - zz_s^\delta(\tau(v') \diamond_s 1)\end{aligned}$$

by (15), Lemma 5.9(i), $x \diamond_t 1 = z_t^\delta$, and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.9(i) and 5.3(ii). If $w = z$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(\tau(v') \diamond_t zz_s^\delta) + z(\tau(v) \diamond_t z_s^\delta) - zz_t^\delta(\tau(v') \diamond_t z_s^\delta) \\ &\qquad - z_t^\delta(\tau(v')y \diamond_t z) + z(\tau(v)y \diamond_t 1) - zz_t^\delta(\tau(v')y \diamond_t 1) \\ &= z_t^\delta\tau(v'y \diamond_s z - v' \diamond_s zz_t^\delta) + z\tau(vy \diamond_s 1 - v \diamond_s z_t^\delta) - zz_t^\delta\tau(v'y \diamond_s 1 - v' \diamond_s z_t^\delta)\end{aligned}$$

by (16) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). If $w = z - z_u^\delta$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(\tau(v') \diamond_t z_u^\delta z_s^\delta) + z_u^\delta(\tau(v) \diamond_u z_s^\delta) - zz_u^\delta(\tau(v') \diamond_u z_s^\delta) \\ &\qquad - \big(z_t^\delta(\tau(v')y \diamond_t z_u^\delta) + z_u^\delta(\tau(v)y \diamond_u 1) - zz_u^\delta(\tau(v')y \diamond_u 1)\big) \\ &= z_t^\delta\tau(v'y \diamond_s (z - z_u^\delta) - v' \diamond_s (z - z_u^\delta)z_t^\delta) + z_u^\delta\tau(vy \diamond_s 1 - v \diamond_s z_u^\delta) - zz_u^\delta\tau(v'y \diamond_s 1 - v' \diamond_s z_u^\delta)\end{aligned}$$

by (15) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). If $w = w'z$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(\tau(v') \diamond_t \tau(w)z_s^\delta) + z(\tau(v) \diamond_t \tau(w')z_s^\delta) - zz_t^\delta(\tau(v') \diamond_t \tau(w')z_s^\delta) \\ &\qquad - \big(z_t^\delta(\tau(v')y \diamond_t \tau(w)) + z(\tau(v)y \diamond_t \tau(w')) - zz_t^\delta(\tau(v')y \diamond_t \tau(w'))\big) \\ &= z_t^\delta\tau(v \diamond_s w - v' \diamond_s wz_t^\delta) + z\tau(vy \diamond_s w' - v \diamond_s w'z_t^\delta) - zz_t^\delta\tau(v \diamond_s w' - v' \diamond_s w'z_t^\delta)\end{aligned}$$

by (16) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). If $w = w'(z - z_u^\delta)$,

$$\begin{aligned}\tau(\text{R.H.S.}) &= z_t^\delta(\tau(v') \diamond_t \tau(w)z_s^\delta) + z_u^\delta(\tau(v) \diamond_u \tau(w')z_s^\delta) - zz_u^\delta(\tau(v') \diamond_u \tau(w')z_s^\delta) \\ &\qquad - \big(z_t^\delta(\tau(v')y \diamond_t \tau(w)) + z_u^\delta(\tau(v)y \diamond_u \tau(w')) - zz_u^\delta(\tau(v')y \diamond_u \tau(w'))\big) \\ &= z_t^\delta\tau(v'y \diamond_s w - v' \diamond_s wz_t^\delta) + z_u^\delta\tau(vy \diamond_s w' - v \diamond_s w'z_u^\delta) - zz_u^\delta\tau(v'y \diamond_s w' - v' \diamond_s w'z_u^\delta)\end{aligned}$$

by (15) and the induction hypothesis. This is τ(L.H.S.) because of Lemmas 5.2 and 5.3(ii). Thus (14) holds and we complete the proof. □

Lemma 5.9. *For $s, t \in \mu_r$, $v, v' \in \mathcal{A}_1$, and $w \in \mathcal{A}_r$, the following equalities hold*:

(i) $vv' \diamond_s 1 = (v \diamond_s 1)(v' \diamond_s 1)$.

(ii) $vy \diamond_s w(z - z_t^\delta) = (vy \diamond_s w - v \diamond_s wz_t^\delta)(y \diamond_t 1)$.

(iii) $yv \diamond_s zw = (y \diamond_s 1)(v \diamond_s zw) + z(yv \diamond_s w) - z(y \diamond_s 1)(v \diamond_s w)$.

(iv) $yv \diamond_s z_t^\delta w = (y \diamond_s 1)(v \diamond_s z_t^\delta w) + z_t^\delta(yv \diamond_t w)$.

(v) $zv \diamond_s zw = z(v \diamond_s zw + zv \diamond_s w - z(v \diamond_s w))$.

(vi) $zv \diamond_s z_t^\delta w = z(v \diamond_s z_t^\delta w) + z_t^\delta(zv \diamond_t w) - zz_t^\delta(v \diamond_t w)$.

(vii) $\tau(v \diamond_s 1) = \tau(v) \diamond_s 1$.

Proof. (i): If $v = 1$ or $v' = 1$, it is obvious. Otherwise, it is enough to show when $v = z^{k_1-1}y\cdots z^{k_m-1}y$ and $v' = z^{l_1-1}y\cdots z^{l_n-1}y$. One calculates

$$vv' \diamond_s 1 = z^{k_1-1}(z - z_s^\delta)\cdots z^{k_m-1}(z - z_s^\delta)z^{l_1-1}(z - z_s^\delta)\cdots z^{l_n-1}(z - z_s^\delta),$$

which is clearly equal to $(v \diamond_s 1)(v' \diamond_s 1)$.

(ii): This is a direct consequence of Lemmas 5.2 and 5.3(ii).

(iii): We first consider the case $v = 1$. If $w = 1$, it is obvious because of Lemma 5.2. If $w = w'z$ ($w' \in \mathcal{A}_r$), the left-hand side turns into

$$(y \diamond_s zw')z = ((y \diamond_s 1)zw' + z(y \diamond_s w') - z(y \diamond_s 1)w')z$$

by Lemma 5.2 and the induction hypothesis on degree of words. This is equal to the right-hand side again by Lemma 5.2. If $w = w'z_t^\delta$ ($w' \in \mathcal{A}_r$, $t \in \mu_r$), the left-hand side turns into

$$(y \diamond_s zw')z_t^\delta + zw(y \diamond_t 1) = (y \diamond_s 1)zw + z(y \diamond_s w')z_t^\delta - z(y \diamond_s 1)w + zw(y \diamond_t 1)$$

by Lemma 5.3(ii) and the induction hypothesis. This is equal to the right-hand side again by Lemma 5.3(ii). If $v = z$, by Lemma 5.2, we have

$$\text{L.H.S.} = (y \diamond_s zw)z$$

and

$$\text{R.H.S.} = ((y \diamond_s 1)(1 \diamond_s zw) + z(y \diamond_s w) - z(y \diamond_s 1)(1 \diamond_s w))z,$$

which are equal as shown just before. If $v = y$, we need to show when $w = 1$, $w'z$, $w'z_t^\delta$ ($w' \in \mathcal{A}_r$, $t \in \mu_r$). If $w = 1$,

$$\text{L.H.S.} = (y^2 \diamond_s 1)z = (y \diamond_s 1)(y \diamond_s z)$$

and

$$\text{R.H.S.} = (y \diamond_s 1)(y \diamond_s z) + z(y^2 \diamond_s 1) - z(y \diamond_s 1)^2,$$

which coincide. If $w = w'z$, by induction on degree of words, the left-hand side turns into

$$(y^2 \diamond_s zw')z = (y \diamond_s 1)(y \diamond_s zw')z + z(y^2 \diamond_s w')z - z(y \diamond_s 1)(y \diamond_s w')z,$$

which is equal to the right-hand side due to Lemma 5.2. If $w = w'z_t^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (y \diamond_s zw)(y \diamond_t 1) + (y^2 \diamond_s zw')z_t^\delta \\ &= ((y \diamond_s 1)zw + z(y \diamond_s w) - z(y \diamond_s 1)w)(y \diamond_t 1) \\ &\qquad + ((y \diamond_s 1)(y \diamond_s zw') + z(y^2 \diamond_s w') - z(y \diamond_s 1)(y \diamond_s w'))z_t^\delta\end{aligned}$$

and

$$\begin{aligned}\text{R.H.S.} = (y \diamond_s 1)(zw(y \diamond_t 1) + (y \diamond_s zw')z_t^\delta) + z((y \diamond_s w)(y \diamond_t 1) + (y^2 \diamond_s w')z_t^\delta) \\ - z(y \diamond_s 1)(w(y \diamond_t 1) + (y \diamond_s w')z_t^\delta),\end{aligned}$$

which are equal. If $v = v'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $v = v'y$, we need to show when $w = 1, w'z, w'z_t^\delta$ $(w' \in \mathcal{A}_r, t \in \mu_r)$. If $w = 1$,

$$\text{L.H.S.} = yv \diamond_s z = (yv \diamond_s 1)z$$

and

$$\text{R.H.S.} = (y \diamond_s 1)(v \diamond_s z) + z(yv \diamond_s 1) - z(y \diamond_s 1)(v \diamond_s 1),$$

which are equal. If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_t^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (yv' \diamond_s zw)(y \diamond_t 1) + (yv \diamond_s zw')z_t^\delta \\ &= ((y \diamond_s 1)(v' \diamond_s zw) + z(yv' \diamond_s w) - z(y \diamond_s 1)(v' \diamond_s w))(y \diamond_t 1) \\ &\qquad + ((y \diamond_s 1)(v \diamond_s zw') + z(yv \diamond_s w') - z(y \diamond_s 1)(v \diamond_s w'))z_t^\delta\end{aligned}$$

and

$$\begin{aligned}\text{R.H.S.} = (y \diamond_s 1)((v' \diamond_s zw)(y \diamond_t 1) + (v \diamond_s zw')z_t^\delta) + z((yv' \diamond_s w)(y \diamond_t 1) + (yv \diamond_s w')z_t^\delta) \\ - z(y \diamond_s 1)((v' \diamond_s w)(y \diamond_t 1) + (v \diamond_s w')z_t^\delta),\end{aligned}$$

which coincide.

(iv): We first consider the case $v = 1$. If $w = 1$, it is obvious because of Lemma 5.3(ii). If $w = w'z$ $(w' \in \mathcal{A}_r)$, the left-hand side turns into

$$(y \diamond_s z_t^\delta w')z = ((y \diamond_s 1)z_t^\delta w' + z_t^\delta (y \diamond_t w'))z$$

by Lemma 5.2 and the induction hypothesis on degree of words. This is equal to the right-hand side again by Lemma 5.2. If $w = w'z_u^\delta$ $(w' \in \mathcal{A}_r, u \in \mu_r)$, the left-hand side turns into

$$(y \diamond_s z_t^\delta w')z_u^\delta + z_t^\delta w(y \diamond_u 1) = (y \diamond_s 1)z_t^\delta w + z_t^\delta (y \diamond_t w')z_u^\delta + z_t^\delta w(y \diamond_u 1)$$

by Lemma 5.3(ii) and the induction hypothesis. This is equal to the right-hand side again by Lemma 5.3(ii). If $v = z$, by Lemma 5.2, we have

$$\text{L.H.S.} = (y \diamond_s z_t^\delta w)z$$

and

$$\text{R.H.S.} = ((y \diamond_s 1)(1 \diamond_s z_t^\delta w) + z_t^\delta (y \diamond_t w))z,$$

which are equal as shown just before. If $v = y$, we need to show when $w = 1, w'z, w'z_u^\delta$ $(w' \in \mathcal{A}_r, u \in \mu_r)$. If $w = 1$,

$$\text{L.H.S.} = (y \diamond_s z_t^\delta)(y \diamond_t 1) + (y^2 \diamond_s 1)z_t^\delta$$

and

$$\text{R.H.S.} = (y \diamond_s 1)(y \diamond_s z_t^\delta) + z_t^\delta (y^2 \diamond_t 1),$$

which are equal because of Lemma 5.3(ii). If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_u^\delta$, by the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (y \diamond_s z_t^\delta w)(y \diamond_u 1) + (y^2 \diamond_s z_t^\delta w')z_u^\delta \\ &= ((y \diamond_s 1)z_t^\delta w + z_t^\delta(y \diamond_t w))(y \diamond_u 1) + ((y \diamond_s 1)(y \diamond_s z_t^\delta w') + z_t^\delta(y^2 \diamond_t w'))z_u^\delta\end{aligned}$$

and

$$\text{R.H.S.} = (y \diamond_s 1)(z_t^\delta w(y \diamond_u 1) + (y \diamond_s z_t^\delta w')z_u^\delta) + z_t^\delta((y \diamond_t w)(y \diamond_u 1) + (y^2 \diamond_t w')z_u^\delta),$$

which coincide. If $v = v'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $v = v'y$, we need to show when $w = 1$, $w'z$, $w'z_u^\delta$ ($w' \in \mathcal{A}_r$, $u \in \mu_r$). If $w = 1$,

$$\text{L.H.S.} = (yv' \diamond_s z_t^\delta)(y \diamond_t 1) + (yv \diamond_s 1)z_t^\delta$$

and

$$\text{R.H.S.} = (y \diamond_s 1)(v \diamond_s z_t^\delta) + z_t^\delta(yv \diamond_t 1),$$

which are equal by Lemma 5.3(ii) and the induction hypothesis. If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_u^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (yv' \diamond_s z_t^\delta w)(y \diamond_u 1) + (yv \diamond_s z_t^\delta w')z_u^\delta \\ &= ((y \diamond_s 1)(v' \diamond_s z_t^\delta w) + z_t^\delta(yv' \diamond_t w))(y \diamond_u 1) + ((y \diamond_s 1)(v \diamond_s z_t^\delta w') + z_t^\delta(yv \diamond_t w'))z_u^\delta\end{aligned}$$

and

$$\text{R.H.S.} = (y \diamond_s 1)((v' \diamond_s z_t^\delta w)(y \diamond_u 1) + (v \diamond_s z_t^\delta w')z_u^\delta) + z_t^\delta((yv' \diamond_t w)(y \diamond_u 1) + (yv \diamond_t w')z_u^\delta),$$

which coincide.

(v): If $v = 1$, by using Lemma 5.2, the right-hand side turns into

$$z^2w + zwz - z^2w = zwz,$$

which is equal to the left-hand side. If $v = z$, we have

$$\text{L.H.S.} = (z \diamond_s zw)z = zwz^2$$

and

$$\text{R.H.S.} = z(zwz + wz^2 - zwz) = zwz^2,$$

which coincide. If $v = y$, we need to show when $w = 1$, $w'z$, $w'z_t^\delta$ ($w' \in \mathcal{A}_r$, $t \in \mu_r$). If $w = 1$,

$$\text{L.H.S.} = zy \diamond_s z = zyz$$

and

$$\text{R.H.S.} = z(y \diamond_s z + zy \diamond_s 1 - z(y \diamond_s 1)) = zyz,$$

which coincide. If $w = w'z$, by induction on degree of words, the left-hand side turns into

$$(zy \diamond_s zw')z = z(y \diamond_s zw' + zy \diamond_s w' - z(y \diamond_s w'))z,$$

which is equal to the right-hand side due to Lemma 5.2. If $w = w'z_t^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\text{L.H.S.} = (z \diamond_s zw)(y \diamond_t 1) + (zy \diamond_s zw')z_t^\delta = zwz(y \diamond_t 1) + z(y \diamond_s zw' + zy \diamond_s w' - z(y \diamond_s w'))z_t^\delta$$

and

$$\text{R.H.S.} = z\big(zw(y \diamond_t 1) + (y \diamond_s zw')z_t^\delta + wz(y \diamond_t 1) + (zy \diamond_s w')z_t^\delta - z(w(y \diamond_t 1) + (y \diamond_s w')z_t^\delta)\big),$$

which are equal.

If $v = v'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $v = v'y$, we need to show when $w = 1$, $w'z$, $w'z_t^\delta$ ($w' \in \mathcal{A}_r$, $t \in \mu_r$). If $w = 1$,

$$\text{L.H.S.} = zv \diamond_s z = zvz$$

and

$$\text{R.H.S.} = z(v \diamond_s z + zv \diamond_s 1 - z(v \diamond_s 1)) = zvz,$$

which coincide. If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_t^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (zv' \diamond_s zw)(y \diamond_t 1) + (zv \diamond_s zw')z_t^\delta \\ &= z(v' \diamond_s zw + zv' \diamond_s w - z(v' \diamond_s w))(y \diamond_t 1) + z(v \diamond_s zw' + zv \diamond_s w' - z(v \diamond_s w'))z_t^\delta\end{aligned}$$

and

$$\begin{aligned}\text{R.H.S.} = z\big((v' \diamond_s zw)(y \diamond_t 1) + (v \diamond_s zw')z_t^\delta + (zv' \diamond_s w)(y \diamond_t 1) \\ + (zv \diamond_s w')z_t^\delta - z((v' \diamond_s w)(y \diamond_t 1) + (v \diamond_s w')z_t^\delta)\big),\end{aligned}$$

which are equal.

(vi): If $v = 1$, by using Lemma 5.2, the right-hand side turns into

$$zz_t^\delta w + z_t^\delta(z \diamond_t w) - zz_t^\delta w = z_t^\delta wz,$$

which is equal to the left-hand side. If $v = z$, we have

$$\text{L.H.S.} = (z \diamond_s z_t^\delta w)z = z_t^\delta wz^2$$

and

$$\text{R.H.S.} = zz_t^\delta wz + z_t^\delta wz^2 - zz_t^\delta wz = z_t^\delta wz^2,$$

which coincide. If $v = y$, we need to show when $w = 1$, $w'z$, $w'z_u^\delta$ ($w' \in \mathcal{A}_r$, $u \in \mu_r$). If $w = 1$,

$$\text{L.H.S.} = zy \diamond_s z_t^\delta = z_t^\delta z(t \diamond_t 1) + (zy \diamond_s 1)z_t^\delta$$

and

$$\begin{aligned}\text{R.H.S.} &= z(y \diamond_s z_t^\delta) + z_t^\delta(zy \diamond_t 1) - zz_t^\delta(y \diamond_t 1)\\ &= z(z_t^\delta(y \diamond_t 1) + (y \diamond_s 1)z_t^\delta) + z_t^\delta z(y \diamond_t 1) - zz_t^\delta(y \diamond_t 1),\end{aligned}$$

which coincide. If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_u^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (z \diamond_s z_t^\delta w)(y \diamond_u 1) + (zy \diamond_s z_t^\delta w')z_u^\delta\\ &= z_t^\delta wz(y \diamond_u 1) + z(y \diamond_s z_t^\delta w')z_u^\delta + z_t^\delta(zy \diamond_t w')z_u^\delta - zz_t^\delta(y \diamond_s w')z_u^\delta\end{aligned}$$

and

$$\text{R.H.S.} = z(z_t^\delta w(y \diamond_u 1) + (y \diamond_s z_t^\delta w')z_u^\delta) + z_t^\delta(wz(y \diamond_u 1) + (zy \diamond_t w')z_u^\delta) - zz_t^\delta(w(y \diamond_u 1) + (y \diamond_t w')z_u^\delta),$$

which are equal. If $v = v'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $v = v'y$, we need to show when $w = 1, w'z, w'z_u^\delta$ ($w' \in \mathcal{A}_r$, $u \in \mu_r$). If $w = 1$, by Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (zv' \diamond_s z_t^\delta)(y \diamond_t 1) + (zv \diamond_s 1)z_t^\delta\\ &= (z(v' \diamond_s z_t^\delta) + z_t^\delta(zv' \diamond_t 1) - zz_t^\delta(v' \diamond_t 1))(y \diamond_t 1) + z(v \diamond_s 1)z_t^\delta\end{aligned}$$

and

$$\begin{aligned}\text{R.H.S.} &= z(v \diamond_s z_t^\delta) + z_t^\delta(zv \diamond_t 1) - zz_t^\delta(v \diamond_t 1)\\ &= z((v' \diamond_s z_t^\delta)(y \diamond_t 1) + (v \diamond_s 1)z_t^\delta) + z_t^\delta z(v' \diamond_t 1)(y \diamond_t 1) - zz_t^\delta(v' \diamond_t 1)(y \diamond_t 1),\end{aligned}$$

which are equal. If $w = w'z$, it is obvious by Lemma 5.2 and the induction hypothesis. If $w = w'z_u^\delta$, by using Lemma 5.3(ii) and the induction hypothesis, one calculates

$$\begin{aligned}\text{L.H.S.} &= (zv' \diamond_s z_t^\delta w)(y \diamond_u 1) + (zv \diamond_s z_t^\delta w')z_u^\delta\\ &= (z(v' \diamond_s z_t^\delta w) + z_t^\delta(zv' \diamond_t w) - zz_t^\delta(v' \diamond_t w))(y \diamond_u 1) + (z(v \diamond_s z_t^\delta w') + z_t^\delta(zv \diamond_t w') - zz_t^\delta(v \diamond_t w'))z_u^\delta\end{aligned}$$

and

$$\begin{aligned}\text{R.H.S.} = z((v' \diamond_s z_t^\delta w)(y \diamond_u 1) + (v \diamond_s z_t^\delta w')z_u^\delta) + z_t^\delta((zv' \diamond_t w)(y \diamond_u 1) + (zv \diamond_t w')z_u^\delta)\\ - zz_t^\delta((v' \diamond_t w)(y \diamond_u 1) + (v \diamond_t w')z_u^\delta),\end{aligned}$$

which are equal.

(vii): If $v = 1$, it is obvious. Otherwise, putting $v = z^{k_1-1}y \cdots z^{k_m-1}y$, one calculates

$$\tau(v \diamond_s 1) = \tau(z^{k_1-1}(z - z_s^\delta) \cdots z^{k_m-1}(z - z_s^\delta)) = z_s^\delta z^{k_m-1} \cdots z_s^\delta z^{k_1-1}$$

and

$$\tau(v) \diamond_s 1 = \psi_s(zx^{k_m-1} \cdots zx^{k_1-1}) = \varphi(z_s x^{k_m-1} \cdots z_s x^{k_1-1}),$$

which are equal. □

6. Proof of Theorem 2.1

We prove that the polynomial F_f defined just before Proposition 5.7 satisfies the theorem. The proof goes by induction on $\deg(f)$ for rooted forests f and $\deg(w)$ for words w. First, we prove the theorem when $f = \bullet$. If $w = 1$, we have

$$\tilde{f}(z_s^\delta) = z_s^\delta(z - z_s^\delta)$$

and

$$z_s^\delta(F_f \diamond_s 1) = z_s^\delta(y \diamond_s 1) = z_s^\delta(z - z_s^\delta),$$

which are equal. Suppose $\deg(w) \geq 1$. If $w = w'z$ ($w' \in \mathcal{A}_r$), by [9, Theorem 2.2(d)], which asserts that R_z and any RTM commute, the induction hypothesis, and Lemma 5.2, we have

$$\tilde{f}(z_s^\delta w'z) = \tilde{f}(z_s^\delta w')z = z_s^\delta(F_f \diamond_s w')z = z_s^\delta(F_f \diamond_s w). \tag{17}$$

If $w = w'z_t^\delta$ ($w' \in \mathcal{A}_r$), we have

$$\tilde{f}(z_s^\delta w' z_t^\delta) = \tilde{f}(z_s^\delta w')z_t^\delta + z_s^\delta w' z_t^\delta(z - z_t^\delta)$$

and, by Lemma 5.3,

$$z_s^\delta(y \diamond_s w' z_t^\delta) = z_s^\delta(y \diamond_s w')z_t^\delta + z_s^\delta w' z_t^\delta(z - z_t^\delta),$$

which are equal by the induction hypothesis.

Next, suppose $\deg(f) \geq 2$. If $f = gh$ ($g, h \neq \mathbb{I}$), we have

$$\tilde{f}(z_s^\delta w) = \tilde{g}\tilde{h}(z_s^\delta w) = \tilde{g}(z_s^\delta(F_h \diamond_s w)) = z_s^\delta(F_g \diamond_s (F_h \diamond_s w)) = z_s^\delta((F_g \diamond_1 F_h) \diamond_s w) = z_s^\delta(F_f \diamond_s w)$$

since $\deg(g), \deg(h) < \deg(f)$ and Lemma 5.5. Let f be a rooted tree and put $f = B_+(g)$. When $w = 1$, we have

$$\tilde{f}(z_s^\delta) = R_{z-z_s^\delta} R_{2z-z_s^\delta} R_{z-z_s^\delta}^{-1} \tilde{g}(z_s^\delta) = R_{z-z_s^\delta} R_{2z-z_s^\delta} R_{z-z_s^\delta}^{-1} z_s^\delta(F_g \diamond_s 1) \tag{18}$$

by the induction hypothesis. Since $\psi_s \varphi R_x = R_{z_s^\delta} \psi_s \varphi$ and $\psi_s \varphi R_y = R_{z-z_s^\delta} \psi_s \varphi$ on $\mathcal{A}_1^1$, we have

$$(18) = z_s^\delta(\psi_s \varphi(R_y R_{x+2y} R_y^{-1}(F_g)) = z_s^\delta(F_f \diamond_s 1). \tag{19}$$

Suppose $\deg(w) \geq 1$. If $w = w'z$ ($w' \in \mathcal{A}_r$), we have (17) again (but this time we consider $\deg(f) \geq 2$). If $w = w'z_t^\delta$ ($w' \in \mathcal{A}$) and $\Delta(f) = \sum_{(f)} f' \otimes f''$, we have

$$\tilde{f}(z_s^\delta w' z_t^\delta) = \sum_{(f)} \widetilde{f'}(z_s^\delta w')\widetilde{f''}(z_t^\delta) = \sum_{(f)} z_s^\delta(F_{f'} \diamond_s w')z_t^\delta(F_{f''} \diamond_t 1)$$

by the induction hypothesis on degree of words and (19). This is equal to $z_s^\delta(F_f \diamond_s w)$ by Proposition 5.7.

Uniqueness of F_f is shown as follows. If $F_f' \in \mathcal{A}_1^1$ also satisfies the theorem, we have

$$(F_f - F_f') \diamond_s w = 0$$

for any $s \in \mu_r$ and any $w \in \mathcal{A}_r$. In particular, putting $w = 1$ we have

$$(F_f - F'_f) \diamond_s 1 = 0,$$

and hence

$$F_f - F'_f = \varphi\psi_s^{-1}(0) = 0. \qquad \square$$

7. Proof of Theorem 2.4

For rooted forests f, we define polynomials $G_f \in \mathcal{A}_1^1$ recursively by

- $G_{\mathbb{I}} = 1$,
- $G_\bullet = -y$,
- $G_t = L_{2x+y}(G_f)$ if $t = B_+(f)$ and $f \neq \mathbb{I}$,
- $G_f = G_g \diamond_1 G_h$ if $f = gh$,

where L_v denotes the left multiplication map by v, i.e., $L_v(w) = vw$ ($v, w \in \mathcal{A}_r$). The subscript of G is extended linearly. In [7], we find that $G_f = F_{S(f)}$.

Lemma 7.1. *For $f \in \mathrm{Aug}(\mathcal{H})$, put $\Delta(f) = \sum_{(f)} f' \otimes f''$. Then we have*

$$\sum_{(f)} F_{f'} \diamond_1 G_{f''} = 0.$$

Proof. See [7, Proposition 4.5]. $\square$

Proof of Theorem 2.4. If $f = \bullet$, the theorem holds since $S(f) = -\bullet$, $G_f = -y$, and Theorem 2.1 for $f = \bullet$. Assume $\deg(f) \geq 2$. If $f = gh$ ($g, h \neq \mathbb{I}$), we have

$$\widetilde{S(f)} = \widetilde{S(gh)} = \widetilde{S(h)S(g)} = \widetilde{S(h)}\widetilde{S(g)} = \widetilde{S(g)}\widetilde{S(h)}$$

because the antipode S is an antiautomorphism, $\widetilde{\cdot}$ is an algebra homomorphism, and RTMs commute with each other. Then, since $\deg(g), \deg(h) < \deg(f)$ and Lemma 5.5, we have

$$\widetilde{S(f)}(z_s^\delta w) = \widetilde{S(g)}(\widetilde{S(h)}(z_s^\delta w)) = z_s^\delta(G_g \diamond_s (G_h \diamond_s w)) = z_s^\delta((G_g \diamond_1 G_h) \diamond_s w) = z_s^\delta(G_f \diamond_s w)).$$

If f is a tree, by letting $\Delta(f) = \sum_{(f)} f' \otimes f''$ and Lemma 7.1, we have

$$z_s^\delta(G_f \diamond_s w)) = -z_s^\delta \sum_{\substack{(f) \\ f' \neq \mathbb{I}}} (F_{f'} \diamond_1 G_{f''}) \diamond_s w. \tag{20}$$

By Lemma 5.5, Theorem 2.1, and the induction hypothesis, we have

$$(20) = -z_s^\delta \sum_{\substack{(f) \\ f' \neq \mathbb{I}}} F_{f'} \diamond_s (G_{f''} \diamond_s w) = -\sum_{\substack{(f) \\ f' \neq \mathbb{I}}} \widetilde{f'}(z_s^\delta(G_{f''} \diamond_s w)) = -\sum_{\substack{(f) \\ f' \neq \mathbb{I}}} \widetilde{f'}(\widetilde{S(f'')}(z_s^\delta w)).$$

Since $\sum_{(f)} f' S(f'') = 0$, we get the theorem. $\square$

8. Proof of Theorem 2.5

Lemma 8.1. *For $f \in \mathrm{Aug}(\mathcal{H})$, we have*

$$F_f = -R_y \tau R_y^{-1}(F_{S(f)}).$$

Proof. See [7, Proposition 5.1]. □

Proof of Theorem 2.5. First, we prove the theorem when $w = z_s^\delta w'(z - z_t^\delta) \in z_s^\delta \mathcal{A}_r(z - z_t^\delta)$. By Theorem 2.4, we have

$$\widetilde{S(f)}(w) = z_s^\delta (F_{S(f)} \diamond_s w'(z - z_t^\delta)).$$

We also have

$$\tau \tilde{f} \tau(w) = \tau \tilde{f}(z_t^\delta \tau(w')(z - z_s^\delta)) = \tau\big(z_t^\delta (F_f \diamond_t \tau(w')(z - z_s^\delta))\big) \tag{21}$$

by Theorem 2.1. Then, by Lemma 8.1, we have

$$(21) = -\tau\big(z_t^\delta (R_y \tau R_y^{-1}(F_{S(f)}) \diamond_t \tau(w')(z - z_s^\delta))\big),$$

which is equal to $z_s^\delta (F_{S(f)} \diamond_s w'(z - z_t^\delta))$ because of Proposition 5.8.

Next, we consider when $w = w'z \in \mathcal{A}_r z$. Since R_z and RTMs commute, we have

$$\widetilde{S(f)}(w) = \widetilde{S(f)}(xw')z$$

and

$$\tau \tilde{f} \tau(w) = \tau \tilde{f} \tau(xw'z) = \tau \tilde{f} \tau(xw')z,$$

which are equal by the induction hypothesis. Similarly, since L_z and RTMs commute, we have the same consequence when $w = zw' \in z\mathcal{A}_r$. □

Acknowledgement

The authors would like to thank the referee for some helpful advice. This work is supported by JSPS KAKENHI Grant Numbers JP19K03434, JP23K03059, and JP22K13897, Grant for Basic Science Research Projects from Sumitomo Foundation, and research funding granted by the University of Kitakyushu.

References

[1] T. Arakawa and M. Kaneko, "On multiple L-values", *J. Math. Soc. Japan* **56**:4 (2004), 967–991. MR Zbl

[2] H. Bachmann and T. Tanaka, "Rooted tree maps and the Kawashima relations for multiple zeta values", *Kyushu J. Math.* **74**:1 (2020), 169–176. MR Zbl

[3] A. Connes and D. Kreimer, "Hopf algebras, renormalization and noncommutative geometry", *Comm. Math. Phys.* **199**:1 (1998), 203–242. MR Zbl

[4] A. Dür, *Möbius functions, incidence algebras and power series representations*, Lecture Notes in Math. **1202**, Springer, 1986. MR Zbl

[5] M. Hirose, H. Murahara, and T. Onozuka, "$\mathbb{Q}$-linear relations of specific families of multiple zeta values and the linear part of Kawashima's relation", *Manuscripta Math.* **164**:3-4 (2021), 455–465. MR Zbl

[6] G. Kawashima and T. Tanaka, "Newton series and extended derivation relations for multiple L-values", preprint, 2008. arXiv 0801.3062

[7] H. Murahara and T. Tanaka, "Algebraic aspects of rooted tree maps", *Ramanujan J.* **60**:1 (2023), 123–139. MR Zbl

[8] T. Tanaka, "Rooted tree maps", *Commun. Number Theory Phys.* **13**:3 (2019), 647–666. MR Zbl

[9] T. Tanaka and N. Wakabayashi, "Rooted tree maps for multiple L-values", *J. Number Theory* **240** (2022), 471–489. MR Zbl

Communicated by Andrew Granville
Received 2022-10-30 Revised 2023-10-08 Accepted 2023-11-27

hmurahara@mathformula.page *Department of Mathematics, The University of Kitakyushu, Fukuoka, Japan*

t.tanaka@cc.kyoto-su.ac.jp *Department of Mathematics, Kyoto Sangyo University, Kyoto, Japan*

wakabayashi@osakac.ac.jp *Center of Physics and Mathematics, Osaka Electro-Communication University, Osaka, Japan*

ALGEBRA AND NUMBER THEORY 18:11 (2024)
https://doi.org/10.2140/ant.2024.18.2027

Terminal orders on arithmetic surfaces

Daniel Chan and Colin Ingalls

The local structure of terminal Brauer classes on arithmetic surfaces was classified (2021), generalising the classification on geometric surfaces (2005). Part of the interest in these classifications is that it enables the minimal model program to be applied to the noncommutative setting of orders on surfaces. We give étale local structure theorems for terminal orders on arithmetic surfaces, at least when the degree is a prime $p > 5$. This generalises the structure theorem given in the geometric case. They can all be explicitly constructed as algebras of matrices over symbols. From this description one sees that such terminal orders all have global dimension two, thus generalising the fact that terminal (commutative) surfaces are smooth and hence homologically regular.

1. Introduction

Given a smooth point on a complex variety of dimension d, the étale local structure is Spec R where $R = \mathbb{C}\{x_1, \ldots, x_d\}$, the algebra of algebraic power series in d variables. This result elegantly captures, in an algebraic fashion, the idea that manifolds are all locally Euclidean. In [Chan and Ingalls 2005], a noncommutative analogue of this result was given for the case of orders on a complex surface. Here we extend the result to arbitrary surfaces, at least under some mild hypotheses.

To set the ambient framework, we briefly recall here the minimal model program for surfaces enriched by a Brauer class as developed in [Chan and Ingalls 2005; 2021]. The minimal model program enriched by a Brauer class relies on a restricted class of log surfaces described below. Although this program uses log surfaces and log surface contractions, the resulting terminal minimal models have a structure theory more akin to terminal commutative surfaces than log surfaces. In other words, they are noncommutative analogues of regular surfaces rather than quotient singularities. In particular, for a terminal order, the log surface is log smooth, i.e., a regular surface with a normal crossing divisor, and the orders over these surfaces have global dimension two, which is equivalent to regular in the commutative case. To set up the program, let X be a normal surface with function field K and Λ a sheaf of maximal $\mathcal{O}_X$-orders on X in a central simple K-algebra D. Let $\beta \in \operatorname{Br} K$ be the Brauer class corresponding to D whose index we assume is prime to residue characteristics of X. Of fundamental importance in the original commutative minimal model program is the canonical divisor, and the key to the noncommutative version is the canonical divisor $K_{X,\beta}$ of β on X defined in terms of the ramification data of β as follows. If Λ is

Chan was supported by the Australian Research Council Discovery Project grant DP220102861. Ingalls was partially supported by a Discovery Grant from the National Science and Engineering Research Council of Canada.
MSC2020: 16H10, 16S38.
Keywords: orders, arithmetic surfaces, minimal model program.

not Azumaya at the generic point of an irreducible divisor $C \subset X$, then we say β *ramifies* along C, and it turns out that we can associate to it the *ramification* $a_C(\beta)$ of β along C, which is an element of the torsion étale cohomology group $H^1(K(C), \mathbb{Q}/\mathbb{Z})$. We define the *ramification index* of β along C to be the order e_C of $a_C(\beta)$. Serre duality theory for the order Λ suggests the definition

$$K_{X,\beta} := K_X + \sum_C \left(1 - \frac{1}{e_C}\right) C,$$

where the sum runs over the finitely many ramification curves (alternatively, one can set $e_C = 1$ when β is unramified along C). With this definition, the notions of *discrepancy*, *terminal*, *canonical* etc. naturally follow as in the original commutative minimal model program. Much of classical commutative surface theory goes through such as resolutions of singularities and Castelnuovo's contraction theorem.

We will be concerned with the étale local theory and so let R be a commutative excellent noetherian normal two-dimensional Hensel local domain with fraction field K. In this context, it will be useful to not only consider maximal orders but, more generally, a normal R-order Λ (see Definition 2.1). To this, we attach a localised Brauer class $(\beta, g_{\mathfrak{p}})$ (in Definition 2.2) where the integers $g_{\mathfrak{p}}$ measure how much Λ deviates from being maximal at the codimension-one prime $\mathfrak{p}$.

When the residue field κ is algebraically closed, terminal localised Brauer classes $(\beta, g_{\mathfrak{p}})$ and the corresponding terminal orders were completely classified in [Chan and Ingalls 2005] and shown to always have global dimension two. In that article, it is shown that

i) R is smooth,

ii) β is zero or has ramification along a normal crossing divisor $C_1 \cup C_2$, and

iii) Λ is maximal except possibly along a curve of multiplicity one when $\beta = 0$, or one of the C_i otherwise.

A complete structure theorem was given for terminal normal orders in this case (see [Chan and Ingalls 2005, Section 2]). If Λ is maximal, then it is isomorphic to a full matrix algebra over a symbol $\Delta = R\langle y, z\rangle/(y^m - u, z^m - v, zy - \zeta yz)$ where ζ is some primitive m-th root of unity and $u, v \in R$ is a regular system of parameters. In general, one obtains a triangular modulo z matrix algebra over such symbols as defined in Definition 2.4.

When κ is a finite field of characteristic prime to the order m of β, it turns out there are more possibilities for terminal localised Brauer classes. If m is a prime > 5, the terminal localised Brauer classes were completely classified in [Chan and Ingalls 2021]. The new possibilities are summed up in Definitions 3.1 and 5.1, but, briefly, when R is regular, there is the additional possibility that β is ramified on a single multiplicity-one curve, and, more interestingly, R can also be a type of Hirzebruch–Jung singularity, in which case $\beta \neq 0$, though it is unramified along codimension-one primes of R. Our main theorem is the following result which generalises the aforementioned structure theory of terminal normal orders to this arithmetic situation. Again, symbols feature significantly but in the more general sense of tensor products

of a $\mathbb{Z}/m$-extension of R with a μ_m-extension and graded components skew commute according to the natural pairing $\mathbb{Z}/m \times \mu_m \to \mu_m$ (see Definition 3.2).

Theorem 1.1. *Let R be an excellent noetherian two-dimensional normal Hensel local domain. Let Λ be a terminal R-order whose degree is a prime, say $m > 5$. If the residue field of R contains a primitive m-th root of unity and has trivial Brauer group, then Λ has global dimension two. In fact, all such Λ are explicitly constructed in Propositions 3.3, 3.5 and Theorem 5.7 as various triangular modulo z matrix algebras over symbols.*

To prove the theorem, we construct explicit examples of normal orders with all the possible terminal localised Brauer classes. In the case when R is regular, this is relatively straight forward. In the singular case, we need to show that our Hirzebruch–Jung singularity, defined to have a minimal resolution whose exceptional locus is a string of projective lines defined over the residue field of R, is actually a cyclic quotient singularity. We prove this in Section 4, and give an explicit construction of the regular cyclic cover which is used in the construction of the corresponding terminal orders. The other step is to show that these explicitly constructed orders are sufficiently nice (in particular, have global dimension two) and that any normal order with the same ramification data has to be Morita equivalent to them. We follow the basic framework of [Chan and Ingalls 2005]. Unfortunately, the use of the Cohen structure theorem in that article is unavailable in this setting, so we give a streamlined method avoiding this tool in Section 2.

2. Uniqueness result for regular almost maximal orders

We review basic definitions of orders, their ramification theory and the relationship with the Brauer group. Since the notion of maximal orders is not stable under étale localisation, we review normal orders as introduced in [Chan and Ingalls 2005]. The main result is a uniqueness-type result for certain normal orders which have global dimension two and are maximal everywhere except possibly on a single irreducible divisor. This was proved over an algebraically closed field using a complicated argument in §2.3 of the same reference. We present a streamlined proof here using the classification of normal orders over discrete valuation rings found in the Appendix.

Let R be a noetherian normal domain and K its field of fractions. Given a central simple K-algebra Q, an *order* A in Q is an R-subalgebra such that A is a finitely generated R-module such that $KA = Q$. Then $K \otimes_R A \simeq Q$ so we sometimes dispense with explicitly mentioning Q and say a finite R-algebra A is an R-order if it is a torsion-free R-module such that $K \otimes_R A$ is a central simple K-algebra. We define the *degree* of A to be $\deg A := \deg K \otimes_R A = \sqrt{\dim_K K \otimes_R A}$.

One ought to think of A as a model of the noncommutative "field" Q in this case. The classical noncommutative analogue of the notion of normality is that the order is *maximal*, that is, if A' is another order in Q containing A, then $A = A'$. Unfortunately, this notion is not stable under étale localisation. When R is two dimensional, the following condition was introduced in [Chan and Ingalls 2005] to remedy this defect, taking its cue from Serre's criterion for normality in the commutative case.

Definition 2.1. Let R be a two-dimensional normal domain. An R-order A is said to be *normal* if

(1) A is a reflexive R-module,

(2) for every height-one prime $\mathfrak{p}$, the localisation $A_\mathfrak{p}$ is *normal* in the sense that its radical is principal as a left and right ideal.

The second condition is thoroughly analysed in the Appendix. Note that maximal orders are normal, and that normal orders are tame.

Let R be a two-dimensional normal domain and A be a normal R-order. Since $K \otimes_R A$ is a central simple K-algebra, it determines a corresponding Brauer class $\beta_A \in \operatorname{Br} K$. Given any codimension-one prime $\mathfrak{p} \lhd R$, with corresponding residue field $\kappa(\mathfrak{p})$, there is a ramification map

$$a_\mathfrak{p} : \operatorname{Br} K \to H^1_{\text{ét}}(\kappa(\mathfrak{p}), \mathbb{Q}/\mathbb{Z}).$$

As noted by Artin and Mumford [1972], this map can be interpreted in terms of orders as follows. First note that $a_\mathfrak{p}(\beta_A)$, being an element of $H^1_{\text{ét}}(\kappa(\mathfrak{p}), \mathbb{Q}/\mathbb{Z})$, is given by a cyclic field extension κ' of $\kappa(\mathfrak{p})$ and a choice of generator σ for the Galois group $\operatorname{Gal}(\kappa'/\kappa(\mathfrak{p}))$. Let J be the radical of $A_\mathfrak{p}$ which is principal, and so is generated by an element, say π. Then $\kappa' = Z(A_\mathfrak{p}/J)$ and σ is the automorphism induced by conjugation by π. Note that $a_\mathfrak{p}(\beta_A) = 0$ means that $A_\mathfrak{p}$ is Azumaya so the collection of nonzero $a_\mathfrak{p}(\beta_A)$ is called the ramification data of A. If $\mathfrak{m} \supset \mathfrak{p}$ is a codimension-two prime, then one can also look at the ramification of the field extension $\kappa'/\kappa(\mathfrak{p})$ at $\mathfrak{m}$, which is referred to as *secondary ramification*. The above results are described in more detail in [Artin and de Jong 2004, §1; Chan 2010, §4; 2011; Grieve and Ingalls 2021, §1] where other phenomena such as the cancellation of secondary ramification data are also explained.

If now B is a normal order contained in the maximal order A above, then from the Appendix, we know that $Z(B_\mathfrak{p}/\operatorname{rad} B_\mathfrak{p}) \simeq \prod_{i=1}^{d} \kappa'$ for some d. Furthermore, conjugation by a generator t of $\operatorname{rad} B_\mathfrak{p}$ permutes the d factors cyclically, and conjugation by t^d reduces to σ. The ramification data of B will thus not only include the $a_\mathfrak{p}(\beta_A)$, but also the integers $g_\mathfrak{p} := d$. Since B is generically Azumaya and thus maximal, we find on varying $\mathfrak{p}$ that all but finitely many of the $g_\mathfrak{p}$ will be one.

Definition 2.2. A *localised Brauer class on* R is a pair $(\beta, g_\mathfrak{p})$ consisting of a Brauer class $\beta \in \operatorname{Br} K$ and a function assigning to each codimension-one prime $\mathfrak{p} \lhd R$ a positive integer $g_\mathfrak{p}$ which equals one for all but finitely many $\mathfrak{p}$. In particular, the localised Brauer class of the normal order B above is $(\beta_A, g_\mathfrak{p})$ in the notation of the previous paragraph.

We now give an instance where the localised Brauer class and R-rank of a normal R-order determines the isomorphism class of the order.

Assumption 2.3. *Suppose now that R is a Hensel local two-dimensional normal domain. Let Δ be a maximal R-order in a division ring and suppose that there exists a normal element $z \in \Delta$ such that*

(1) *the quotient $\Delta/z\Delta$ is supported, as an R-module, on a codimension-one prime $\mathfrak{q}$,*

(2) *the element z generates the radical of $\Delta_\mathfrak{q}$,*

(3) *$\Delta/z\Delta$ is hereditary.*

Under these assumptions, we construct the order

$$\Delta_d = \Delta_d(z) := \begin{pmatrix} \Delta & \Delta & \cdots & \Delta \\ z\Delta & \Delta & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ z\Delta & \cdots & z\Delta & \Delta \end{pmatrix} \subseteq M_d(\Delta) \tag{2-1}$$

Definition 2.4. We will refer to the subalgebra Δ_d in (2-1) above as a *triangular modulo z matrix algebra*.

Proposition 2.5. *Under Assumption 2.3, the order Δ_d is normal and has global dimension two. Its localised Brauer class $(\beta, g_{\mathfrak{p}})$ is given by*

(1) *β is the Brauer class of Δ,*

(2) *$g_{\mathfrak{q}} = d$ and all other $g_{\mathfrak{p}} = 1$.*

Proof. Note (1) follows from the fact that Δ_d is an order in $M_n(K\Delta)$. To check Δ_d is a reflexive R-module, note first that the Δ is reflexive being a maximal order. Also, Δ is a domain so z must be a non-zero-divisor. Thus $z\Delta$ and hence also Δ_d are reflexive as well.

We consider now local structure at a codimension-one prime $\mathfrak{p}$. If $\mathfrak{p} \neq \mathfrak{q}$, then from Assumption 2.3(1), we know that $(\Delta_d)_{\mathfrak{p}} \simeq M_d(\Delta_{\mathfrak{p}})$ so is maximal. On the other hand, Assumption 2.3(2) ensures that $(\Delta_d)_{\mathfrak{q}}$ is normal and $g_{\mathfrak{q}} = d$. This completes the verification of (2).

Finally, consider the normal element

$$t := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \vdots & & \ddots & 1 \\ z & 0 & \cdots & \cdots & 0 \end{pmatrix} \in \Delta_d.$$

By Assumption 2.3(3), we see that $\Delta_d/t\Delta_d \simeq \prod_{i=1}^{d} \Delta/z\Delta$ is hereditary, so Δ_d itself must have global dimension two. □

In the light of this proposition, we might consider Δ_d to be regular and almost maximal as in the title of this section.

Theorem 2.6. *Suppose that Assumption 2.3 holds and let Δ_d be the normal R-order of (2-1). Let Λ be any normal R-order with the same localised Brauer class as Δ_d. Then $n = \frac{\deg \Lambda}{\deg \Delta_d}$ is an integer and $\Lambda \simeq M_n(\Delta_d)$ so has global dimension two.*

Proof. Suppose $\deg \Delta_d \mid \deg \Lambda$ and let $\Lambda' := M_n(\Delta_d)$. Since the localised Brauer classes of Λ, Λ' coincide, as do their degrees, we may embed them both in a common central simple K-algebra Q. By Corollary A.6, we know that $\Lambda_{\mathfrak{q}} \simeq \Lambda'_{\mathfrak{q}}$ so by altering one of the embeddings, we may suppose that we actually have $\Lambda_{\mathfrak{q}} = \Lambda'_{\mathfrak{q}}$. Consider the (Λ, Λ')-bimodule $B := (\Lambda\Lambda')^{**} \subset Q$, the reflexive hull of the

R-module $\Lambda\Lambda'$. Now Λ' has global dimension two by Proposition 2.5 and B is Cohen–Macaulay as an R-module so is projective as a Λ'-module by [Ramras 1969, Proposition 3.5] (the hypotheses are stated differently there but the proof applies in our case).

We first show an isomorphism of right Λ' modules, $B_{\Lambda'} \simeq \Lambda'_{\Lambda'}$. Now R is Henselian so Krull–Schmidt holds for Λ'-modules. The indecomposable projective Λ'-modules are isomorphic to summands of Λ' and there are exactly d isomorphism classes of these, say $P_1, \dots, P_d$, corresponding to the rows of Δ_d. From Proposition A.3 there exists an integer r such that $(P_i)_{\mathfrak{q}}/(P_i)_{\mathfrak{q}}(\operatorname{rad}\Lambda'_{\mathfrak{q}}) \simeq S_i^{\oplus r}$ where S_i is a simple $\Lambda'_{\mathfrak{q}}$ module and that, furthermore, the S_i are all nonisomorphic. In particular, two finitely generated projective Λ'-modules are isomorphic if and only if their localisations at $\mathfrak{q}$ are isomorphic. Now by our choice of embeddings, $B_{\mathfrak{q}} = \Lambda'_{\mathfrak{q}}$ so $B \simeq \Lambda'$ as desired.

To complete the proof of the theorem when $\deg \Delta_d \mid \deg \Lambda$, it suffices to show that the natural map $\Lambda \to \operatorname{End}_{\Lambda'} B = \Lambda'$ is an isomorphism. Since both sides are reflexive, this can be checked on codimension-one primes $\mathfrak{p}$. When $\mathfrak{p} \neq \mathfrak{q}$, it is an isomorphism since Λ is maximal. When $\mathfrak{p} = \mathfrak{q}$, it is an isomorphism since we recalibrated so $\Lambda_{\mathfrak{q}} = \Lambda'_{\mathfrak{q}} = B_{\mathfrak{q}}$.

If $\deg \Delta_d$ does not divide $\deg \Lambda$, we apply the special case proved to $M_{\deg \Delta_d}(\Lambda)$, which shows that at least Λ is Morita equivalent to Δ_d. Hence $\Lambda \simeq \operatorname{End}_{\Delta_d} P$ for some projective Δ_d-module. We can argue as before, looking locally at $\mathfrak{q}$ to see that d indecomposable projective modules occur equally in the decomposition of P, so $P \simeq \Delta_d^{\oplus n}$ for some n as desired. □

3. Toral terminal orders, regular centre case

We have a series of concepts that are motivated by toric or toroidal geometry, but that does not satisfy either of these definitions. We call these *toral*. Let $(R, \mathfrak{m})$ be a two-dimensional noetherian regular Hensel local domain with field of fractions K. We introduce the notion of a *toral terminal* localised Brauer class on R. These are all terminal. In the special setting of [Chan and Ingalls 2021], i.e., when R is an arithmetic surface with finite residue field and the ramification data are all p-torsion for some prime p, this exactly agrees with terminal. Assuming that R has enough roots of unity and a trivial Brauer group, we then classify the normal R-orders associated to toral terminal localised Brauer classes and show they all are regular in the sense that they have global dimension two.

In the regular centre case, toral terminal localised Brauer classes fall into two types. The first, without secondary ramification is defined below.

Definition 3.1. A localised Brauer class $(\beta, g_{\mathfrak{p}})$ on R is *toral terminal without secondary ramification* if there exists a regular system of parameters $u, v \in R$ such that

(1) β is unramified at every codimension-one prime $\mathfrak{p}$ except possibly $\mathfrak{p} = (u)$ (that is, $a_{\mathfrak{p}}(\beta) = 0$ for $\mathfrak{p} \neq (u)$), and

(2) all $g_{\mathfrak{p}} = 1$ except possibly $\mathfrak{p} = (v)$.

We will construct orders via symbols and so need to assume the existence of enough roots of unity.

Definition 3.2. Suppose $\zeta \in R$ is a primitive n-th root of unity. Given $a, b \in R - 0$ we define the *R-symbol* $(a, b) := (a, b)_\zeta^R$ to be the R-algebra

$$\Lambda = \frac{R\langle x, y\rangle}{(x^n - a,\ y^n - b,\ yx - \zeta xy)}.$$

Proposition 3.3. *Let R be a two-dimensional regular noetherian Hensel local domain and $(\beta, g_\mathfrak{p})$ be a toral terminal localised Brauer class with ramification as given in Definition 3.1. Suppose that* $\operatorname{Br} R/\mathfrak{m} = 0$ *and R possesses a primitive n-th root of unity ζ where n is the order of β in the* $\operatorname{Br} K$.

Let $a \in R$ be any element chosen so the ramification of β along (u) is given by adjoining an n-th root of a. Then

(1) *$\Delta = (u, a)_\zeta^R$ is a maximal order in a division ring with the same ramification data as β,*

(2) *any normal order Λ with localised Brauer class $(\beta, g_\mathfrak{p})$ is isomorphic to $M_m(\Delta_d(v))$ (see notation in (2-1)) where $d = g_{(v)}$ and m is an arbitrary positive integer.*

In particular, Λ has global dimension two.

Proof. As secondary ramification must cancel, the ramification of β along (u) must be given by an étale cyclic extension of $R/(u)$, say of degree n. Now the cyclic étale extensions of $R/\mathfrak{m}$, $R/(u)$ and R all coincide, so we may find $a \in R$ defining this ramification, by adjoining $\sqrt[n]{a}$. Also, since $\operatorname{Br} R = \operatorname{Br} R/\mathfrak{m} = 0$, we know n is the order of β.

Note that $\Delta = (u, a)_\zeta^R$ is Azumaya on the open set $u \neq 0$ and is maximal at the generic point $\mathfrak{p}$ of $u = 0$ and has the same ramification as β at $\mathfrak{p}$. Now reflexive orders which are maximal in codimension one are maximal globally by [Auslander and Goldman 1960, Theorem 1.5]. Thus since Δ is also reflexive, it is a maximal order which has the same ramification as that of β. Furthermore, as already observed, $\operatorname{Br} R = 0$ so both β and Δ determine the same Brauer class in $\operatorname{Br} K$.

We seek now to apply Theorem 2.6. We begin by verifying Assumption 2.3 for $z = v$. First, Δ is a domain since its degree coincides with the period. Clearly $\Delta/v\Delta$ is supported along the prime (v) only, and in fact v generates the radical of the localisation $\Delta_{(v)}$. It remains to verify Assumption 2.3(3). Let $x \in \Delta$ be the n-th root of u as in Definition 3.2. Then x gives a non-zero-divisor in $\Delta/v\Delta$. Furthermore, $\Delta/(x, v)$ is the separable extension $(R/\mathfrak{m})\big(\sqrt[n]{a}\big)$ so $\Delta/v\Delta$ is indeed hereditary. □

Definition 3.4. A localised Brauer class $(\beta, g_\mathfrak{p})$ on R is *toral terminal with secondary ramification* if there exists a regular system of parameters such that

(1) β is unramified away from (uv),

(2) β is ramified along both (u) and (v) and the ramification at these prime ideals are given by totally ramified field extensions of the residue fields,

(3) all $g_\mathfrak{p}$ equal 1 except possibly $g_{(v)}$.

More generally (but still assuming R regular), we say a localised Brauer class $(\beta, g_{\mathfrak{p}})$ is *toral terminal* if it is either toral terminal with secondary ramification as above, or toral terminal without secondary ramification as in Definition 3.1.

Proposition 3.5. *Let R be a two-dimensional regular noetherian Hensel local domain and $(\beta, g_{\mathfrak{p}})$ be a toral terminal localised Brauer class with secondary ramification as given in Definition 3.4. Suppose that $\operatorname{Br} R/\mathfrak{m} = 0$ and R possesses a primitive n-th root of unity where n is the order of β in $\operatorname{Br} K$. Any normal order Λ with localised Brauer class $(\beta, g_{\mathfrak{p}})$ is isomorphic to $M_m(\Delta_d(y))$ (see notation in (2-1)) where*

(1) $d = g_{(v)}$,

(2) $\Delta_d(y)$ *is built from the maximal order*

$$\Delta = (au, bv)^R_\zeta,$$

where $a, b \in R^\times$ are units, ζ is an appropriate n-th root of unity and y is the n-th root of bv used in Definition 3.2.

In particular, Λ has global dimension two.

Proof. We use Theorem 2.6 along the same lines as the proof of Proposition 3.3. It suffices to find u, v, ζ such that $(u, v)^R_\zeta$ has the same ramification as β.

Let κ_u denote the residue field at the point (u). The ramification $a_{(u)}(\beta) \in H^1(\kappa_u, \mathbb{Q}/\mathbb{Z})$ of β along (u) corresponds to a cyclic field extension $\tilde{\kappa}/\kappa_u$ and a generator of the Galois group. Since we assumed existence of primitive n-th roots of unity, we may use Kummer theory to see that $\tilde{\kappa} = \kappa_u\big(\sqrt[n]{\bar{v}}\big)$ for some $\bar{v} \in \kappa_u$. Since this is a totally ramified extension, we may change generators and assume that $\bar{v}$ is the restriction of bv for some $b \in R^\times$. If σ is the chosen generator of the Galois group, then $\sigma\big(\sqrt[n]{bv}\big) = \zeta\sqrt[n]{bv}$ where ζ is the n-th root of unity required in (2) above. Arguing the same way for ramification of β along (v) and using the fact that secondary ramification cancels, we see that $a_{(v)}(\beta)$ is given by adjoining an n-th root of au for some $a \in R^\times$ and we are done. □

4. Hirzebruch–Jung singularities as cyclic quotient singularities

Over the complex numbers the Hirzebruch–Jung singularities are well understood and all arise as cyclic quotient singularities. We show a similar result for two-dimensional normal singularities which are Hirzebruch–Jung in the sense that their minimal resolution is a string of projective lines defined over the residue field. A related result can be found in [Kollár 2013, Theorem 3.32], the difference being that we do not assume the existence of an underlying ground field.

More precisely, suppose that R is a two-dimensional normal noetherian excellent commutative Hensel local domain with residue field κ. Suppose it is a *κ-rational Hirzebruch–Jung* singularity in the sense that it has a rational minimal resolution $f : Y \to \operatorname{Spec} R =: X$ such that the exceptional locus is a string $E_1, \dots, E_r$ of exceptional curves isomorphic to the projective line over κ, that is, all E_i are isomorphic to $\mathbb{P}^1_\kappa$, E_i intersects E_{i+1} in a single point which is κ-rational and there are no other intersections. There

are more complicated analogues of Hirzebruch–Jung singularities studied in the literature which we have not studied as they do not arise in the study of the terminal orders considered in this paper.

Since R is Hensel local, we may choose irreducible curves $E_0, E_{r+1} \subset Y$ such that E_0 (respectively, E_{r+1}) intersects E_1 (respectively, E_r) in a single κ-rational point and E_0 and E_2 (respectively, E_{r+1} and E_{r-1}) are disjoint. Let $m_i = -E_i^2$ which, by minimality of the resolution f, must be at least two. We define pairs $v_0 = (0, 1)$, $v_1 = (1, 0)$ and then recursively define

$$v_{i+1} = m_i v_i - v_{i-1} \quad \text{for } i = 1, \ldots, r. \tag{4-1}$$

When R is defined over the complex numbers, these give the exceptional curves in the toric description of the Hirzebruch–Jung singularity. An easy induction shows that the dot product $(1, 1) \cdot v_i$ weakly increases with i, so $(m, -k) := v_{r+1}$ satisfies $0 < k < m$. Similarly, one can show that k, m are relatively prime. The integer m appears in our key theorem below.

Theorem 4.1. *Let R be a κ-rational Hirzebruch–Jung singularity as defined above and $X = \operatorname{Spec} R$. The Weil divisor f_*E_0 is m-torsion in the sense that mf_*E_0 is Cartier. There is a natural ring structure on*

$$S := \bigoplus_{l=0}^{m-1} \mathcal{O}_X(-lf_*E_0)t^l$$

making it a regular local ring with the same residue field κ as R. In particular, if R contains m-th roots of unity, then R is a cyclic quotient singularity. Furthermore, S/R is étale away from the singular point.

Before launching in to the proof, we set up the appropriate theory first. The idea is to recover as much of the toric theory of Hirzebruch–Jung singularities over the complex numbers as possible. To this end we consider:

Definition 4.2. A divisor D on Y is *toral* if it belongs to $\bigoplus_{i=0}^{r+1} \mathbb{Z}E_i$. We say $w \in R$ is *toral* or is a *toral function* if the associated divisor of f^*w is toral.

Remark 4.3. Definition 4.2 depends on the choice of E_0 and E_{r+1}.

Our first order of business is to classify all toral functions and show their divisors on Y are given by the lattice points in the cone $\mathbb{R}_{\geq 0}(0, 1) + \mathbb{R}_{\geq 0}(m, -k)$ where we recall $(m, -k) = v_{r+1}$. Any function $w \in R$ is determined, up to $H^0(\mathcal{O}_Y^\times) = R^\times$, by its divisor (f^*w) on Y. It thus suffices to classify effective toral divisors D on Y such that $D \sim 0$. Now f is a rational resolution, so by [Lipman 1969, Proposition 11.1 i)], we know $D \sim 0$ if and only if $D.E_i = 0$ for $i = 1, \ldots, r$.

To enumerate all such toral divisors, we work as follows. Let $L = \bigoplus_{i=0}^{r+1} \mathbb{Z}E_i$ and consider $L^* := \operatorname{Hom}_{\mathbb{Z}}(L, \mathbb{Z}) = \bigoplus \mathbb{Z}E_i^*$ where $\{E_i^*\}$ is a dual basis to the E_i. For $i = 1, \ldots, r$, let $E_i^\vee := (C \mapsto E_i.C) \in L^*$ and $\mathbb{E} < L^*$ be the subgroup generated by the $E_i^\vee$. The next result follows from (4-1).

Proposition 4.4. *The homomorphism $v : L^* \to \mathbb{Z}^2$ defined by $E_i^* \mapsto v_i$ is surjective with kernel $\mathbb{E}$.*

The $D \sim 0$ condition ensures that the naturally induced map $L^* \to \mathbb{Z}$ given by $\chi \mapsto \chi(D)$ actually factors through ν to give a homomorphism $\lambda_D : \mathbb{Z}^2 \to \mathbb{Z}$. Note that

$$D = \sum_{i=0}^{r+1} \lambda_D(\nu_i) E_i. \tag{4-2}$$

Suppose now that λ_D is given by dot product with $(i, j) \in \mathbb{Q}^2$. Now $\nu_0 = (0, 1)$, $\nu_1 = (1, 0)$ so $i, j \in \mathbb{Z}$. The divisor D corresponding to (i, j) is effective when $(i, j) \cdot \nu_0 \geq 0$, $(i, j) \cdot \nu_{r+1} \geq 0$, that is, (i, j) lies in the cone $\mathbb{R}_{\geq 0}(1, 0) + \mathbb{R}_{\geq 0}(k, m)$. We now abuse notation and write $x^i y^j$ for any toral function with this D as its divisor in Y. The notation allows us to write $x^i y^j x^{i'} y^{j'} = x^{i+i'} y^{j+j'}$ with the caveat that it holds only modulo $R^\times$. We summarise the results up to this point.

Proposition 4.5. *The toral functions are $x^i y^j$ where $(i, j) \in \mathbb{R}_{\geq 0}(1, 0) + \mathbb{R}_{\geq 0}(k, m)$. Its divisor in Y is $\sum \lambda_l E_l$ where $\lambda_l = (i, j) \cdot \nu_l$.*

Corollary 4.6. *The Weil divisor $f_* E_0$ is m-torsion.*

Proof. The divisor of zeros $\sum \lambda_l E_l$ of the toral function $x^k y^m$ has $\lambda_0 = m$, $\lambda_{r+1} = 0$. □

Let $\mathfrak{m}$ be the maximal ideal of R and Z be the fundamental cycle so $\mathfrak{m}^n = f_* \mathcal{O}(-nZ)$ for any positive integer n [Liu 2002, Lemma 9.4.14]. Note that $Z = E_1 + \cdots + E_r$. Consider a Weil divisor $C \subset \operatorname{Spec} R$ so $\mathcal{O}(-C)$ is a reflexive ideal. Note that $f^* \mathcal{O}(-C)/\mathcal{T} \simeq \mathcal{O}_Y(-\widetilde{C})$ for some $\mathfrak{m}$-torsion sheaf $\mathcal{T}$ and Cartier divisor $\widetilde{C}$ which is the strict transform of C away from the exceptional locus. Here, by $\mathfrak{m}$-torsion, we mean that $\mathcal{T}$ is annihilated by some positive power of $\mathfrak{m}$.

Definition 4.7. We call $\widetilde{C}$ the *pullback* of C. This agrees with the usual pullback in the case that C is Cartier.

Reflexivity ensures that $f_* \mathcal{O}_Y(-\widetilde{C}) = \mathcal{O}(-C)$ so $\mathcal{O}(-C)$ is *contracted* in the language of [Lipman 1969, Definition 6.1]. Suppose more generally that $D \in \operatorname{Div} Y$ is such that $\mathcal{O}_Y(-D)$ is generated by global sections so [Lipman 1969, Corollary to 7.3] ensures that $\mathfrak{m}^n f_* \mathcal{O}_Y(-D) = f_* \mathcal{O}_Y(-D - nZ)$. Applying f_* to the exact sequence

$$0 \to \mathcal{O}_Y(-D - Z) \to \mathcal{O}_Y(-D) \to \mathcal{O}_Z(-D) \to 0$$

gives the exact sequence

$$0 \to f_* \mathcal{O}(-D) \otimes_R R/\mathfrak{m} \to H^0(\mathcal{O}_Z(-D)) \to R^1 f_* \mathcal{O}_Y(-D - Z).$$

Now $\mathcal{O}_Y(-D - Z)$ is generated by global sections since the same is true of $\mathcal{O}_Y(-D)$ and $\mathcal{O}_Y(-Z)$, so $R^1 f_* \mathcal{O}_Y(-D - Z)) = 0$ from which follows the next result.

Lemma 4.8. *Let D be a divisor on Y such that $\mathcal{O}_Y(-D)$ is generated by global sections. Write $I = f_* \mathcal{O}_Y(-D)$ which will be an ideal in R if D is effective. Then we have $I \otimes_R R/\mathfrak{m} \simeq H^0(\mathcal{O}_Z(-D))$. In particular, a set of generators for I can be found by giving a set of global sections of $\mathcal{O}_Y(-D)$ whose*

restriction to Z gives a spanning set for $H^0(\mathcal{O}_Z(-D))$. This applies in particular to $I = \mathcal{O}(-C)$ where C is an effective Weil divisor on X and $D = \widetilde{C}$ is the pullback.

We will use this lemma to find toral generators for $\mathcal{O}(-if_*E_0)$. First, we need a "toral" basis for $H^0(Z, \mathcal{L})$ where $\mathcal{L}$ is a line bundle on Z with all $d_i := \deg_{E_i} \mathcal{L}$ greater than or equal to 0.

Definition 4.9. The intersections $E_i \cap E_{i+1}$, $i = 0, \ldots, r$, are said to be the *toral points* of Z. A nonzero section $s \in H^0(Z, \mathcal{L})$ is *toral* if its zero set is a union of exceptional curves and toral points. We say s is *basic toral* if it also satisfies the following condition: whenever $s|_{E_i} \neq 0$ but has a zero at $E_i \cap E_{i+1}$ (respectively, $E_i \cap E_{i-1}$), then $s|_{E_j} = 0$ for $j > i$ (respectively, $j < i$).

We can construct basic toral sections as follows. Start with some nonzero section $s_i \in H^0(E_i, \mathcal{L}|_{E_i}) \simeq H^0(\mathbb{P}^1, \mathcal{O}(d_i))$ which is "toral" in the sense that its zeros are confined to $E_{i-1} \cup E_{i+1}$. Up to a scalar in κ, there are $d_i + 1$ of these. We show it can be extended uniquely to a basic toral section s of $\mathcal{L}$. Now if s_i has a zero at $E_{i-1} \cap E_i$, then we simply extend by setting $s|_{E_j} = 0$ for $j < i$. If on the other hand s_i is nonzero at $E_{i-1} \cap E_i$, then there is a unique way to extend it to a toral section on E_{i-1} and we can continue by induction. A similar argument determines s on E_j for $j > i$. This gives the following:

Lemma 4.10. *Any basic toral section s is uniquely determined by any nonzero restriction $s|_{E_l}$ and has the form constructed in the preceding paragraph.*

Proposition 4.11. *Given a line bundle $\mathcal{L}$ on Z with nonnegative degrees $d_i := \deg_{E_i} \mathcal{L} \geq 0$, there exists a basis for $H^0(Z, \mathcal{L})$ consisting of basic toral sections. This basis is unique up to scaling the basis elements.*

Proof. First, $H^0(Z, \mathcal{L})$ is naturally isomorphic to the kernel of the natural map $\bigoplus_{l=1}^{r} H^0(E_l, \mathcal{O}_{E_i}(d_i)) \to H^0(T, \mathcal{O}_T)$ where T is the set of nodes in Z. This has dimension $d = \sum(d_i + 1) - (r - 1)$. The only linear relations between basic toral sections are those which are scalar multiples of each other so it suffices to find d basic toral sections, no two of which are multiples of each other. The above construction provides these once we note that the basic toral section constructed from some toral section $s_i \in H^0(E_i, \mathcal{L}|_{E_i})$ coincides with one constructed from $s_{i-1} \in H^0(E_{i-1}, \mathcal{L}_{E_{i-1}})$ if and only if s_i, s_{i-1} take on the same nonzero value at $E_{i-1} \cap E_i$. □

We can now construct toral generators for reflexive ideals in R.

Proposition 4.12. *Let D be an effective toral divisor on Y such that $\mathcal{O}_Y(-D)$ is generated by global sections and $I = f_*\mathcal{O}_Y(-D)$ be the associated ideal of R. Then I is generated by toral functions.*

Proof. Combining Lemma 4.8 with Proposition 4.11, it suffices to lift every basic toral section of $\mathcal{O}_Z(-D)$ to a toral section of $\mathcal{O}_Y(-D)$. Let $d_i := -D.E_i$ and consider a basic toral section whose restriction to E_i has a zero of order e at E_{i-1} and order $d_i - e$ at E_{i+1}. Lifting this to a toral function amounts to finding an effective toral divisor $\Delta = \sum_{j=0}^{r+1} \delta_j E_j$ such that

a) $-D \sim \Delta$, and

b) $\delta_{i-1} = e$, $\delta_i = 0$, $\delta_{i+1} = d_i - e$.

Condition a) amounts to checking that all the intersection numbers $(D+\Delta).E_j$ equal 0. We solve these equations for δ_j by induction on $|j-i|$ and simultaneously prove effectivity of Δ by proving δ_j is nondecreasing for $j \geq i$ and nonincreasing for $j \leq i$. The base case is satisfied since

$$(D+\Delta).E_i = \delta_{i-1} + D.E_i + \delta_{i+1} = e - d_i + (d_i - e) = 0.$$

Suppose now that nondecreasing integers $\delta_i, \ldots, \delta_j$ have now been defined satisfying $(D+\Delta).E_l = 0$ for $l = i, \ldots, j$. We examine the equation

$$0 = (D+\Delta).E_j = -d_j + \delta_{j-1} - m_j\delta_j + \delta_{j+1}.$$

We may thus solve for the integer δ_{j+1} which further satisfies

$$\delta_{j+1} - \delta_j = d_j + ((m_j - 1)\delta_j - \delta_{j-1}) \geq 0$$

by the inductive hypothesis. A similar argument works for nonincreasing δ_j when $j \leq i$. □

Proof. We can now complete the proof of Theorem 4.1. Inspired by Corollary 4.6, or rather its proof, we define a ring structure on

$$S = \bigoplus_{l=0}^{m-1} \mathcal{O}(-lf_*E_0)t^l$$

by defining $t^{-m} = x^k y^m$. Given a toral function $x^i y^j \in \mathcal{O}(-lf_*E_0)$, we say

$$x^i y^j t^l = x^{i-kl/m} y^{j-l}$$

is a toral function in S. It suffices to find two toral functions in $f_1,\ f_2 \in S$ which generate the maximal ideal

$$\mathfrak{n} := \mathfrak{m} \oplus \mathcal{O}(-f_*E_0)t \oplus \cdots \oplus \mathcal{O}(-(m-1)f_*E_0)t^{m-1}.$$

By Proposition 4.12, it suffices to show $f_1,\ f_2$ will generate all the toral elements in the summands $\mathfrak{m}, \ldots, \mathcal{O}(-(m-1)f_*E_0)t^{m-1}$. Our sloppiness in notation for $x^i y^j$ is warranted since we only care about the ideal generated by $f_1,\ f_2$.

From Proposition 4.5, we know that $x^i y^j$ is a toral function in $\mathcal{O}(-lf_*E_0)$ if and only if $(i,j) \in \mathbb{R}_{\geq 0}(1,0) + \mathbb{R}_{\geq 0}(k,m)$ and furthermore $l \leq (i,j).v_0 = j$. We may thus let

$$f_1 = x^k y^m t^{m-1} = x^{k/m} y. \tag{4-3}$$

To find f_2 we first find $l \in \{1, \ldots, m-1\}$ which solves $kl \equiv -1 \bmod m$, which is possible since k and m are relatively prime. Let $i = \frac{kl+1}{m}$ and

$$f_2 = x^i y^l t^l = x^{1/m}. \tag{4-4}$$

An elementary calculation shows that the toral elements of S all have the form $x^i y^j$ where $i \in \frac{1}{m}\mathbb{Z}$, $j \in \mathbb{Z}$ and $0 \leq j \leq \frac{m}{k} i$. It follows that all the toral elements in $\mathfrak{n}$ are generated by f_1 and f_2.

Finally, the construction of S here is the cyclic covering trick, see, for example, [Lazarsfeld 2004, 4.1.B], which away from the singularity uses an m-torsion line bundle so S/R is étale. □

For use in the next section, we record the following fact which follows from Proposition 4.5.

Lemma 4.13. *The toral function $f_1 \in S$ defined in (4-3) is such that $f_1^m \in R$ and its divisor is mf_*E_0.*

5. Toral terminal orders, singular centre case

Unlike in the geometric case where the residue fields are algebraically closed, there are now terminal orders with singular centre [Chan and Ingalls 2021]. Their ramification data were classified in the case where the "index" [Chan and Ingalls 2021, §3, p. 6] was a prime $m > 5$. Here we classify the corresponding orders, giving explicit constructions of them.

Throughout this section, we let $(R, \mathfrak{m})$ be an excellent normal two-dimensional noetherian Hensel local domain with residue field κ. The classification of terminal ramification data on R is best encapsulated via the following definition.

Definition 5.1. A localised Brauer class $(\beta, g_{\mathfrak{p}})$ on R is *toral terminal* if either R is regular and we are in the case of Definitions 3.1 or 3.4, or if the following hold:

(1) R is a κ-rational Hirzebruch–Jung singularity whose residue field has trivial Brauer group. Let $E_1, \dots, E_r$ be the string of exceptional curves in the minimal resolution (indexed naturally so E_i intersects E_{i+1}).

(2) β is unramified along codimension-one primes in R.

(3) The order m of β equals the *determinant* of R which is defined to be $\det R := \det(M_R)$ where

$$M_R := -\begin{pmatrix} E_1^2 & 1 & 0 & \cdots & 0 \\ 1 & E_2^2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & E_r^2 \end{pmatrix}.$$

(4) At most one $g_{\mathfrak{p}}$ does not equal 1, in which case $\mathfrak{p}$ corresponds to an irreducible curve C on the minimal resolution which (scheme-theoretically) intersects the exceptional curve in a single κ-rational point $y \notin E_2 \cup \cdots \cup E_{r-1}$.

Remark 5.2. (1) Given a toral terminal localised Brauer class $(\beta, g_{\mathfrak{p}})$ on R as above, [Chan and Ingalls 2021, Theorem 7.1] shows that it is terminal whenever m is a prime > 2. If m is a prime > 5 and κ is finite, then these are the only terminal localised Brauer classes on singular R.

(2) In the same article, the residue fields are all assumed to be finite and so have trivial Brauer group.

Let $(R, \mathfrak{m})$ be a κ-rational Hirzebruch–Jung singularity of determinant m (see Definition 5.1(3)). Suppose that κ contains a primitive m-th root of unity. From Theorem 4.1, there exists a regular local

ring $(S, \mathfrak{n})$ such that $S/\mathfrak{n} = \kappa$ and S/R is a cyclic extension which is étale away from the singular point. In Theorem 4.1, it is presented as a $\mathbb{Z}/m$-graded algebra $S = \bigoplus_{i \in \mathbb{Z}/m} S_i$ so $(\mathbb{Z}/m)^\vee = \mu_m$ acts on it naturally. Let $\alpha \in H^1(\kappa, \mathbb{Z}/p)$ correspond to a cyclic degree-m field extension $\tilde{\kappa}/\kappa$ with some chosen action of $\mathbb{Z}/m$. Since R is Hensel local, there is a corresponding cyclic étale extension $\tilde{R}/R$ and a corresponding μ_m-graded decomposition $\tilde{R} = \bigoplus_{\omega \in \mu_m} \tilde{R}_\omega$.

Definition 5.3. We define the *symbol* (S, α) to be the R-algebra whose underlying R-module structure is given by

$$\Delta := S \otimes_R \tilde{R}$$

and multiplication given by the skew-commutation relations

$$rs = \omega^i sr \quad \text{for all } s \in S_i,\ r \in \tilde{R}_\omega.$$

Proposition 5.4. *The symbol $\Delta = (S, \alpha)$ defined above is a maximal order in a division ring, and Δ is Azumaya in codimension one.*

Proof. The commutative algebra $S \otimes_R \tilde{R}$ is an étale extension of S and hence regular and thus Cohen–Macaulay. It follows that Δ is a reflexive R-module. In codimension one, both S/R and $\tilde{R}/R$ are étale so Δ is defined using the usual symbol construction of Azumaya algebras. We see thus that Δ is a maximal order and is Azumaya in codimension one.

It only remains to show that $\Delta_K := \Delta \otimes_R K(R)$ is a division ring, which we do by showing that it cannot have period $< m$. Suppose that $\tilde{R}$ is obtained by adjoining an m-th root of $\alpha \in R^\times$ to R. Suppose that the order of Δ_K is $n \mid m$. If σ denotes the action of some fixed primitive m-th root of unity on S, then the cyclic algebra $A := K(S)[z; \sigma]/(z^m - \alpha^n)$ is a full matrix algebra over $K(R)$. We see thus from [Gille and Szamuely 2006, Corollary 4.7.5] that $\alpha^n \in K(R)$ is a norm from $K(S)$, say $\alpha^n = N(\beta)$, where $\beta = \beta_1 \beta_2^{-1}$ for $\beta_1, \beta_2 \in S$. Now S is a UFD so we may prime factorise both β_1 and β_2.

We first show that by modifying β by a 1-coboundary $\gamma^{-1}\sigma(\gamma)$, $\gamma \in K(S)$, we may assume that $\beta \in S^\times$. Note that $N(\beta_1)$, $N(\beta_2)$ differ by the unit $\alpha^n \in R^\times$. Thus if there is any prime factor $p_2 \mid \beta_2$, there is some prime factor $p_1 \mid \beta_1$ and i such that $p_2 \mid \sigma^i(p_1)$. We may thus multiply by some 1-coboundary so that these factors now cancel. Having reduced the number of prime factors of β_2, we are done by induction.

We now use the fact that $S/\mathfrak{n} = R/\mathfrak{m}$ to see that modulo $\mathfrak{m}$, α^n is a m-th power. Since $\tilde{\kappa}$ is a degree-m extension of κ obtained by adjoining an m-th root of α, we must have $n = m$. □

Proposition 5.5. *Let $(\beta, g_\mathfrak{p})$ be a toral terminal localised Brauer class on a κ-rational Hirzebruch–Jung singularity R. Suppose that R has a primitive m-th root of unity where m is the order of β. Then there is an $\alpha \in H^1(\kappa, \mathbb{Z}/m)$ such that class of the symbol (S, α) in $\operatorname{Br} K(R)$ is β.*

Proof. We use the notation in Definition 5.1. Let $X \to \operatorname{Spec} R$ be the minimal resolution of the κ-rational Hirzebruch–Jung singularity R. Note that β is ramified only along the exceptional curve, so by the Artin–Mumford–Saltman sequence [1972; 2008, Theorem 6.12], the ramification covers of the E_i are all étale. Now $\operatorname{Br}\kappa = 0$ so the ramification along E_i is given by cyclic cover of the form $\mathbb{P}^1_{\kappa_i} \to \mathbb{P}^1_\kappa \simeq E_i$

where κ_i is a cyclic extension of κ of degree $n \mid m$. The ramification is thus given by an element of $H^1(\kappa, \mathbb{Z}/p)$.

We now use the theory developed in [Chan and Ingalls 2021, Setion 4]. There was an assumption there that the residue field was finite, but the theory goes through in this case, as long as one realises that the absolute Galois group G of κ is now not necessarily $\hat{\mathbb{Z}}$. In particular, we have the following version of [Chan and Ingalls 2021, Proposition 9.8].

Lemma 5.6. *There exists a homomorphism $z : \mathbb{Z}^2 \to H^1(\kappa, \mathbb{Z}/m)$ such that the ramification of β along E_i is given by $z(v_i)$ where v_i is as defined in (4-1).*

In particular, we see that the ramification of β, and hence β itself is completely determined by $z(0, 1) = z(v_0) = 0$ and $z(1, 0) = z(v_1)$, the ramification along E_1. It is now clear that we can pick $\alpha \in H^1(\kappa, \mathbb{Z}/m)$ so that the symbol (S, α) has the same ramification as β along E_1, and hence belongs to the same Brauer class over $K(R)$. □

Theorem 5.7. *Let Λ be a normal order over an excellent two-dimensional Hensel local noetherian domain $(R, \mathfrak{m})$ which is not regular. Suppose its localised Brauer class $(\beta, g_{\mathfrak{p}})$ is toral terminal. If R has a primitive m-th root of unity where m is the order of β, then $\Lambda \simeq M_n(\Delta_d(z))$ where*

(1) *Δ is the symbol (S, α) where S is the regular cyclic cover of R constructed in Theorem 4.1 and $\alpha \in H^1(\kappa, \mathbb{Z}/m)$,*

(2) *$n \in \mathbb{N}$, d is either 1 or the unique $g_{\mathfrak{p}}$ not equal to 1, and $z \in S \subset \Delta$ is the normal element denoted f_1 in (4-3).*

In particular, Λ has global dimension two.

Proof. From Proposition 5.5, we may choose α so that $\Delta = (S, \alpha)$ represents the Brauer class β. We also know from Proposition 2.5 that Δ is a maximal order in a division ring. The result will thus follow from Theorem 2.6 once we verify Assumption 2.3. Let $f : X \to \operatorname{Spec} R$ be the minimal resolution. Using the notation in Section 4, toral terminal implies that we may pick $E_0 \subset X$ to be such that $C := f_* E_0$ corresponds to the codimension-one prime $\mathfrak{q}$ with $g_{\mathfrak{q}} \neq 1$ if such a prime exists (and is otherwise an arbitrary prime divisor intersecting $E_1 \setminus E_2$ in a κ-rational point).

Note that z is a toral function and hence gives a normal element of Δ. We also know from Lemma 4.13 that $z^m \in R$ and that its associated divisor is $m f_* E_0$. It follows that $(z) \lhd S$ is the unique prime lying over $\mathfrak{q}$. Thus $\Delta/z\Delta$ is supported on C as an R-module and Assumption 2.3(1) is verified. It also follows that z lies in the radical of $\Delta_{\mathfrak{q}}$. Consider now

$$\bar{\Delta} := \Delta/z\Delta \simeq S/(z) \otimes_{R/\mathfrak{q}} \tilde{R},$$

where $\tilde{R}$ is the cyclic étale extension of R determined by α. To see that z generates the radical of $\Delta_{\mathfrak{q}}$ it suffices to observe that $\bar{\Delta}_{\mathfrak{q}}$ is a central simple $K(R/\mathfrak{q})$-algebra since it is readily identified with a symbol. This completes the verification of Assumption 2.3(2) so it remains only to show that $\bar{\Delta}$ is hereditary. To this end, let y be the other generator of $\operatorname{rad} S$ denoted f_2 in (4-4). It is normal in Δ and thus $\bar{\Delta}$. Then

$\bar{\Delta}/y\bar{\Delta} \simeq \kappa \otimes_R \tilde{R}$ which is a field and hence has global dimension zero. This completes the proof of the theorem. □

Corollary 5.8. *Let Λ be a normal order over an excellent two-dimensional normal noetherian Hensel local domain R with finite residue field. Suppose that its localised Brauer class $(\beta, g_{\mathfrak{p}})$ is terminal,*

(1) *the order m of β is prime > 5, and*

(2) *R has primitive m-th roots of unity.*

Then Λ has global dimension two.

Appendix

The theory of normal and more generally hereditary orders over a complete discrete valuation ring is well known and can be found in standard texts such as Reiner's classic text [1975]. We extend some results to arbitrary discrete valuation rings R which are not necessarily complete.

Let $\mathfrak{m}$ be the maximal ideal of R and K be its field of fractions. Let Δ be a maximal order in some K-central division ring $K\Delta$. Let $\Lambda \subseteq M_n(\Delta)$ be a hereditary order in $M_n(K\Delta)$ with say Jacobson radical J. Note that $\Delta/\operatorname{rad}\Delta$ is central simple, say isomorphic to $M_r(D)$ where D is a division ring. The case when R is complete is simpler because we always have $r = 1$ then.

Proposition A.1. *Consider a right projective Λ-module P such that* $\operatorname{End} P \simeq \Delta$. *Then $P/PJ \simeq S^{\oplus r}$ where S is a simple Λ-module and r is the integer such that $\Delta/\operatorname{rad}\Delta \simeq M_r(D)$ for some division ring D. This result holds in particular for $P = \Delta^n$.*

Proof. Consider the natural ring homomorphism

$$\Delta = \operatorname{End}_\Lambda P \to \operatorname{End}_\Lambda P/PJ.$$

This map is surjective since P is projective. From the ideal theory of maximal orders [Auslander and Goldman 1960, Theorem 2.3], the only semisimple quotient of Δ is $\Delta/\operatorname{rad}\Delta \simeq M_r(D)$ so P/PJ must be the direct sum of r copies of a single simple.

Finally, $\operatorname{End}_\Lambda \Delta^n = \operatorname{End}_{M_n(\Delta)} \Delta^n$ since Λ is an order in $M_n(K\Delta)$. The final statement follows thus from $\Delta = \operatorname{End}_{M_n(\Delta)} \Delta^n$ which is a consequence of Morita theory. □

Definition A.2. We say that Λ is *normal* if its Jacobson radical J is free of rank 1 as a left and right module.

Suppose from now on that Λ is normal, so that one can choose a *uniformiser* $t \in J$ such that $J = \Lambda t = t\Lambda$ so the inner automorphism $r \mapsto trt^{-1}$ induces an automorphism of the semisimple ring Λ/J. We refer to this as the *t-action* on the Wedderburn components, which induces an analogous t-action on the simples (it maps a simple S to $S \otimes_\Lambda t\Lambda$). Suppose there are d simples $S_1, \ldots, S_d$. The following show that the t-action permutes the Wedderburn components cyclically.

Proposition A.3. *The action of t permutes all the simples cyclically. In other words, by reindexing if necessary, we may assume that $S_i^{\oplus r} \simeq Pt^{i-1}/Pt^i$ where $P = \Delta^n$.*

Proof. From Proposition A.1, we may assume S_1 is the simple such that $P/Pt \simeq S_1^{\oplus r}$. The composition factors of any finite-length quotient of P all lie in the t-orbit of S_1. For any simple S_i, we may choose a finitely generated projective Λ-module Q which surjects onto S_i. If the $M_n(K\Delta)$-module $Q \otimes_R K$ is isomorphic to $(K\Delta^n)^{\oplus a}$, then by clearing denominators, we can find an embedding $Q \hookrightarrow P^{\oplus a}$ such that the cokernel of $Qt \hookrightarrow P^{\oplus a}$ has finite length. It follows that Q/Qt has composition factors in the t-orbit of S_1 so, in particular, S_i lies in the t-orbit. □

For our structure theory, we will need the order

$$\Delta_d := \begin{pmatrix} \Delta & \Delta & \cdots & \Delta \\ \operatorname{rad}\Delta & \Delta & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \operatorname{rad}\Delta & \cdots & \operatorname{rad}\Delta & \Delta \end{pmatrix} \subseteq M_d(\Delta),$$

whose radical is

$$\operatorname{rad}\Delta_d := \begin{pmatrix} \operatorname{rad}\Delta & \Delta & \cdots & \Delta \\ \operatorname{rad}\Delta & \operatorname{rad}\Delta & & \vdots \\ \vdots & \ddots & \ddots & \Delta \\ \operatorname{rad}\Delta & \cdots & \operatorname{rad}\Delta & \operatorname{rad}\Delta \end{pmatrix}.$$

If $\pi \in \Delta$ is a generator for $\operatorname{rad}\Delta$, then one readily shows that

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ \pi & 0 & \cdots & \cdots & 0 \end{pmatrix} \in \Delta_d$$

generates the radical on the left and right so Δ_d is normal.

The order Δ_d arises naturally:

Proposition A.4. *Let P be the projective Λ-module Δ^n.*

(1) *As subsets of* $\operatorname{Hom}_{M_n(K\Delta)}(P \otimes_R K, P \otimes_R K) = K\Delta$, *we have*

$$\operatorname{Hom}_\Lambda(Pt^i, Pt^j) = \begin{cases} \Delta & \text{if } 0 \le i-j < d, \\ \operatorname{rad}\Delta & \text{if } 0 < j-i \le d. \end{cases}$$

(2) *In particular*, $\operatorname{End}_\Lambda(P \oplus Pt \oplus \cdots \oplus Pt^{d-1}) = \Delta_d$.

Proof. Part (2) follows from (1) which we now prove. We may as well assume that $i = 0$. First, $\operatorname{Hom}_\Lambda(P, Pt)$ is the kernel of the natural surjection $\operatorname{End}_\Lambda P \to \operatorname{End}_\Lambda P/Pt$ which in turn is $\operatorname{rad}\Delta$. For

$-d < j \leq 0$, the composition factors of Pt^j/P are all nonisomorphic to the simple summands of P/Pt. Hence $\mathrm{Hom}_\Lambda(P, Pt^j) = \mathrm{Hom}_\Lambda(P, P) = \Delta$ in this case. A similar argument shows that

$$\mathrm{Hom}_\Lambda(P, Pt) = \mathrm{Hom}_\Lambda(P, Pt^2) = \cdots = \mathrm{Hom}_\Lambda(P, Pt^d). \qquad \square$$

Theorem A.5. *Let R be a normal order over a discrete valuation ring R, and r be the integer in Proposition A.1. Then there exists $b \mid r$ such that $M_b(\Lambda) \simeq M_c(\Delta_d)$ for some c.*

Proof. We know from Proposition A.3 that conjugation by t permutes the Wedderburn components of $\Lambda/\mathrm{rad}\,\Lambda$ so Λ_Λ must be the projective cover of a semisimple module of the form $(S_1 \oplus \cdots \oplus S_d)^{\oplus a}$ for some a. We let $c = l/r$, $b = l/a$ where l is the lowest common multiple of a and r. It follows that $\Lambda^{\oplus b} \simeq (P \oplus Pt \oplus \cdots \oplus Pt^{d-1})^{\oplus c}$. Using Proposition A.4 to compute endomorphism rings of both sides gives the theorem. $\square$

Corollary A.6. *Up to isomorphism, a normal R-order Λ is uniquely determined by the Brauer class of $K\Delta$ (in $\mathrm{Br}\,K$), the number of simples of Λ and the degree (or R-rank) of Λ.*

Proof. Suppose the Morita equivalence between Λ and Δ_d is given by the Morita bimodule ${}_\Lambda Q_{\Delta_d}$. If S is the direct sum of one copy of each simple Δ_d-module, then Q_{Δ_d} is the projective cover of a semisimple module $S^{\oplus a}$ for some a. The degree of Λ determines a uniquely. $\square$

To determine Λ itself, it suffices to classify all possible indecomposable projective Δ_d-modules Q and their tops $Q/Q(\mathrm{rad}\,\Delta_d)$, a task which we address now.

Let $\bar{\Delta} = \Delta/\mathrm{rad}\,\Delta$. We define a $\bar{\Delta}$-*flag* to be a sequence of $\bar{\Delta}$-submodules

$$0 \leq \bar{I}_1 \leq \bar{I}_2 \leq \cdots \leq \bar{I}_d = \bar{\Delta}.$$

Their inverse images in Δ gives the sequence of Δ-modules

$$\mathrm{rad}\,\Delta \leq I_1 \leq I_2 \leq \cdots \leq I_d = \Delta.$$

The module of row vectors $Q = (I_1\ I_2\ \cdots\ I_d)$ defines a Δ_d-submodule of Δ^d. It is projective since Δ_d is normal.

Proposition A.7. *The order Δ_d is normal.*

(1) *The projective Δ_d-module Q constructed from a $\bar{\Delta}$-flag as above is indecomposable.*

(2) $Q/Q(\mathrm{rad}\,\Delta_d) \simeq (\bar{I}_1\ I_2/I_1 \cdots I_d/I_{d-1})$.

(3) *Every indecomposable projective Δ_d-module has this form.*

(4) *In particular, the indecomposable projectives are precisely the projective covers of any direct sum of r simple Δ_d-modules.*

Proof. Note that $Q \otimes_R K \simeq (K\Delta)^d$ is an indecomposable $M_d(K\Delta)$-module so Q is also indecomposable.

If S denotes a simple Δ-module, then the simple Δ_d-modules are

$$(S\ 0 \cdots 0), (0\ S\ 0 \cdots 0), \ldots, (0 \cdots 0\ S).$$

Thus Q is the projective cover of the semisimple module

$$Q/Q(\operatorname{rad}\Delta_d) \simeq (\bar{I}_1\ I_2/I_1 \cdots I_d/I_{d-1}),$$

which is a direct sum of exactly r simples. By varying the $\bar{\Delta}$-flag, we can construct the projective cover of any direct sum of r simples we like. It thus remains to show there are no other indecomposable projective modules. Let L be one such and suppose $L/L(\operatorname{rad}\Delta_d)$ is a direct sum of more than r simples. Then we can find a direct summand T consisting of precisely r simples and use an appropriate $\bar{\Delta}$-flag to construct the projective cover Q of T. Then the natural surjection $L \to T$ lifts to a surjection $L \to Q$ which must split, a contradiction. If on the other hand, $L/L(\operatorname{rad}\Delta_d)$ had fewer than r simples, then we could apply the same argument to show that some projective Q constructed using a $\bar{\Delta}$-flag decomposes, another contradiction. □

Example A.8. In the case where $d = r$, we can get examples of normal orders which differ most significantly from $\Delta_d = \Delta_r$. Let Q be the indecomposable projective Δ_r-module corresponding to the "complete" $\bar{\Delta}$-flag where all I_{i+1}/I_i are simple. The top $Q/Q(\operatorname{rad}\Delta_r)$ of Q contains exactly one copy of every simple Δ_r-module so the resulting order $\Lambda := \operatorname{End}_{\Delta_r} Q$ will indeed be normal. Interestingly, Λ is an order in $K\Delta$ so is much smaller than Δ_r.

References

[Artin and de Jong 2004] M. Artin and A. J. de Jong, "Stable orders over surfaces", preprint, 2004, available at http://www.math.lsa.umich.edu/courses/711/ordersms-num.pdf.

[Artin and Mumford 1972] M. Artin and D. Mumford, "Some elementary examples of unirational varieties which are not rational", *Proc. Lond. Math. Soc.* (3) **25** (1972), 75–95. MR Zbl

[Auslander and Goldman 1960] M. Auslander and O. Goldman, "Maximal orders", *Trans. Amer. Math. Soc.* **97** (1960), 1–24. MR Zbl

[Chan 2010] K. Chan, *Resolving singularities of orders over surfaces*, Ph.D. thesis, University of New South Wales, 2010.

[Chan 2011] D. Chan, "Lectures on orders", lecture notes, 2011, available at https://web.maths.unsw.edu.au/~danielch/Lect_Orders.pdf.

[Chan and Ingalls 2005] D. Chan and C. Ingalls, "The minimal model program for orders over surfaces", *Invent. Math.* **161**:2 (2005), 427–452. MR Zbl

[Chan and Ingalls 2021] D. Chan and C. Ingalls, "The minimal model program for arithmetic surfaces enriched by a Brauer class", preprint, 2021. arXiv 2108.03105

[Gille and Szamuely 2006] P. Gille and T. Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Stud. Adv. Math. **101**, Cambridge Univ. Press, 2006. MR Zbl

[Grieve and Ingalls 2021] N. Grieve and C. Ingalls, "On the Kodaira dimension of maximal orders", *Adv. Math.* **392** (2021), art. id. 108013. MR Zbl

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Math. **200**, Cambridge Univ. Press, 2013. MR Zbl

[Lazarsfeld 2004] R. Lazarsfeld, "Vanishing theorems", Chapter 4, pp. 239–267 in *Positivity in algebraic geometry*, *I*: *Classical setting*: *line bundles and linear series*, Ergebnisse der Math. (3) **48**, Springer, 2004.

[Lipman 1969] J. Lipman, "Rational singularities, with applications to algebraic surfaces and unique factorization", *Inst. Hautes Études Sci. Publ. Math.* **36** (1969), 195–279. MR Zbl

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Grad. Texts in Math. **6**, Oxford Univ. Press, 2002. MR Zbl

[Ramras 1969] M. Ramras, "Maximal orders over regular local rings of dimension two", *Trans. Amer. Math. Soc.* **142** (1969), 457–479. MR Zbl

[Reiner 1975] I. Reiner, *Maximal orders*, Lond. Math. Soc. Monogr. **5**, Academic Press, London, 1975. MR Zbl

[Saltman 2008] D. J. Saltman, "Division algebras over surfaces", *J. Algebra* **320**:4 (2008), 1543–1585. MR Zbl

Communicated by Jason P. Bell
Received 2022-11-24 Revised 2023-07-05 Accepted 2023-11-27

danielc@unsw.edu.au *School of Mathematics and Statistics, University of New South Wales Sydney, Sydney, Australia*

cingalls@math.carleton.ca *School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada*

ALGEBRA AND NUMBER THEORY 18:11 (2024)
https://doi.org/10.2140/ant.2024.18.2047

Word measures on $\mathrm{GL}_N(q)$ and free group algebras

Danielle Ernst-West, Doron Puder and Matan Seidel

Fix a finite field K of order q and a word w in a free group $\boldsymbol{F}$ on r generators. A w-random element in $\mathrm{GL}_N(K)$ is obtained by sampling r independent uniformly random elements $g_1, \dots, g_r \in \mathrm{GL}_N(K)$ and evaluating $w(g_1, \dots, g_r)$. Consider $\mathbb{E}_w[\mathrm{fix}]$, the average number of vectors in K^N fixed by a w-random element. We show that $\mathbb{E}_w[\mathrm{fix}]$ is a rational function in q^N. If $w = u^d$ with u a nonpower, then the limit $\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}]$ depends only on d and not on u. These two phenomena generalize to all stable characters of the groups $\{\mathrm{GL}_N(K)\}_N$.

A main feature of this work is the connection we establish between word measures on $\mathrm{GL}_N(K)$ and the free group algebra $K[\boldsymbol{F}]$. A classical result of Cohn (1964) and Lewin (1969) is that every one-sided ideal of $K[\boldsymbol{F}]$ is a free $K[\boldsymbol{F}]$-module with a well-defined rank. We show that for w a nonpower, $\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{C}{q^N} + O\big(\frac{1}{q^{2N}}\big)$, where C is the number of rank-2 right ideals $I \le K[\boldsymbol{F}]$ which contain $w-1$ but not as a basis element. We describe a full conjectural picture generalizing this result, featuring a new invariant we call the q-primitivity rank of w.

In the process, we prove several new results about free group algebras. For example, we show that if T is any finite subtree of the Cayley graph of $\boldsymbol{F}$, and $I \le K[\boldsymbol{F}]$ is a right ideal with a generating set supported on T, then I admits a basis supported on T. We also prove an analog of Kaplansky's unit conjecture for certain $K[\boldsymbol{F}]$-modules.

1. Introduction

Fix $r \in \mathbb{Z}_{\ge 1}$. We let $\boldsymbol{F}$ denote the free group on r generators. A word $w \in \boldsymbol{F}$ induces a map on any finite group, $w : G^r \to G$, by substituting the letters of w with elements of G. This map defines a distribution on the group G: the pushforward of the uniform distribution on G^r. Equivalently, this distribution is

MSC2020: primary 20C07, 20E05; secondary 16S34, 20C33, 20G40, 20H30, 68R15.
Keywords: free group algebra, word measures, q-primitivity rank.

the normalized number of times each element in G is obtained by a substitution in w. We call such a distribution a *word measure* on G, and if w is given, *the w-measure on G*. For example, if $w = abab^{-2}$, a w-random element in G is $ghgh^{-2}$ where g, h are independent, uniformly random elements of G.

The study of word measures on various families of groups revealed structural depth with surprising connections to objects in combinatorial and geometric group theory (see, e.g., [Puder 2014; Puder and Parzanchevski 2015; Magee and Puder 2019; 2021; 2024; Hanany and Puder 2023]). It has proven useful for many questions regarding free groups and their automorphism groups (see, e.g., [Puder and Parzanchevski 2015; Hanany et al. 2020]), as well as for questions about random Schreier graphs and their expansion (see, e.g., [Puder 2015; Hanany and Puder 2023]). Previous works in the subject study word measures on the groups $\mathrm{Sym}(N)$, $U(N)$, $O(N)$, Sp(N) and generalized symmetric groups. Section 1E explains how some of the results in the current paper relate to the established structure in other families of groups.

We focus on word measures on $\mathrm{GL}_N(K)$, the general linear group over a fixed finite field K of order q. As seen in other families of groups, word measures on this family demonstrate structural depth. Most interestingly, we show that the analysis of word measures on $\mathrm{GL}_N(K)$ is intertwined with the theory of free group algebras.

1A. ***The average number of fixed vectors.*** We consider various families of real- or complex-valued functions defined on $\mathrm{GL}_N(K)$, and study their expected value under word measures. Our core example is the function $\mathrm{fix} : \mathrm{GL}_N(K) \to \mathbb{Z}_{\geq 0}$ counting elements in the vector space $V = K^N$ which are fixed by a given matrix in $\mathrm{GL}_N(K)$. Not only does this special case illustrate our more general results, but is also a case in which our understanding goes deeper. The function fix is, in fact, a family of functions, one for every value of $N \in \mathbb{Z}_{\geq 1}$. We let $\mathbb{E}_w[\mathrm{fix}]$ denote the expected value of fix under the w-measure on $\mathrm{GL}_N(K)$, so $\mathbb{E}_w[\mathrm{fix}]$ is also a sequence of numbers, one for every value of $N \in \mathbb{Z}_{\geq 1}$. Our first result is the following.

Theorem 1.1. *For every $w \in \boldsymbol{F}$ and every large enough N, $\mathbb{E}_w[\mathrm{fix}]$ is given by a rational function in q^N with rational coefficients.*

For example, if $w = [a, b] = aba^{-1}b^{-1}$ is the commutator of two basis elements, then

$$\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{(q-1)^2 q^N - (q-1)^3}{(q^N - 1)(q^N - q)}$$

for every $N \geq 2$ (recall that $q = |K|$ is fixed throughout, so this expression is indeed a rational function in q^N with coefficients in $\mathbb{Q}$). Consult Table 1 for further examples. For general words, the rational expression is valid for every $N \geq |w|$. See Section 2 for a tighter lower bound on N. Theorem 1.1 is a special case of Theorem 1.11.

Our second result alludes to a result of Nica [1994]. Let $1 \neq w = u^d$ where $d \in \mathbb{N}_{\geq 1}$ and u a nonpower. Nica proved, inter alia, that the distribution of the number of fixed points in a w-random permutation in $\mathrm{Sym}(N)$ has a limit distribution as $N \to \infty$ which depends solely on d and not on u. A similar phenomenon was later shown to hold in various other families of groups. We add the groups $\{\mathrm{GL}_N(K)\}_N$

w	q	$\mathbb{E}_w[\mathrm{fix}]$	valid for
a	every q	2	$N \geq 1$
a^2	q even	3	$N \geq 2$
	q odd	4	
a^3	$q \equiv 0, 2 \pmod 3$	4	$N \geq 3$
	$q \equiv 1 \pmod 3$	8	
$[a, b]$	every q	$2 + \frac{(q-1)^2 q^N - (q-1)^3}{(q^N-1)(q^N-q)}$	$N \geq 2$
a^2b^3	$q = 2$	$2 + \frac{2}{2^N-2}$	$N \geq 3$
	$q = 3$	$2 + \frac{4}{3^N-3}$	
$[a, b]^2$	$q = 2$	$3 + \frac{2(2^{2N}-9\cdot 2^N+26)}{(2^N-1)(2^N-2)(2^N-8)}$	$N \geq 4$
$a^2b^2c^2$	$q = 2$	$2 + \frac{1}{(2^N-2)^2}$	$N \geq 2$
	$q = 3$	$2 + \frac{8(3^{2N}-4\cdot 3^N+5)}{(3^N-1)^2(3^N-3)^2}$	

Table 1. The rational expressions giving $\mathbb{E}_w[\mathrm{fix}]$ for various words $w \in \boldsymbol{F}(a, b, c)$ and various values of $q = |K|$. For the first four words, rational expressions are given for all values of q. For the remaining three words, rational expressions are given only for particular values of q.

as such a family. In our illustrative special case, this is captured by the following result, which first appeared in [Ernst-West 2019]. It also appeared independently in [Eberhard and Jezernik 2022, Section 8].

Theorem 1.2. *Let $1 \neq w = u^d$ with $d \geq 1$ and u a nonpower. Then*

$$\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}] = \#\{p \in K[x] : p \mid x^d - 1 \text{ and } p \text{ monic}\}. \tag{1-1}$$

In particular, the limit does not depend on u.

Combined with Theorem 1.1, if c_d is the number of monic divisors of $x^d - 1 \in K[x]$, we get that $\mathbb{E}_w[\mathrm{fix}] = c_d + O\big(\frac{1}{q^N}\big)$. In particular, for nonpowers, $\mathbb{E}_w[\mathrm{fix}] = 2 + O\big(\frac{1}{q^N}\big)$, and for proper powers $c_d \geq 3$ (if $d \geq 2$, then $x^d - 1$ admits at least three distinct monic divisors: 1, $x - 1$ and $x^d - 1$). Theorem 1.2 is analogous to the result in the symmetric group $\mathrm{Sym}(N)$, where this limit is equal to the number of positive divisors of d in $\mathbb{Z}$ [Nica 1994]. In fact, it is sufficient to prove that the limit in (1-1) depends only on d and not on u, and then the left-hand side of (1-1) is equal to $\lim_{N\to\infty} \mathbb{E}_{a^d}[\mathrm{fix}]$. This number can then be extracted from the analysis of uniformly random elements in $\mathrm{GL}_N(K)$ — see, for example, [Fulman and Stanton 2016]. Theorem 1.2 is a special case of the more general Theorem 1.12 below.

1B. ***The q-primitivity rank.*** The analysis of $\mathbb{E}_w[\mathrm{fix}]$, yielding Theorems 1.1 and 1.2, can be performed using elementary linear algebraic arguments. In fact, this is how they were first derived in the [Ernst-West

2019]. However, it turns out to be extremely useful to analyze these quantities using the theory of free group algebras.

Denote by $\mathcal{A} \stackrel{\text{def}}{=} K[\boldsymbol{F}]$ the free group algebra over K: its elements are finite linear combinations of elements of the free group $\boldsymbol{F}$ with coefficients from the finite field K. It is a classical result of Cohn [1964] and Lewin[1] [1969] that right ideals of $\mathcal{A}$ are free right $\mathcal{A}$-modules with a well-defined rank.[2] An analogous result holds for left ideals, but here we use right ideals only — in fact, from now on, we write "ideals" to mean "right ideals". In Section 2 below we derive a formula for $\mathbb{E}_w[\mathrm{fix}]$ as a sum over a finite set of finitely generated ideals of $\mathcal{A}$, and Section 3 shows that the contribution of every such ideal is of order determined by its rank.

In particular, this algebraic perspective allows a further understanding of the deviation of $\mathbb{E}_w[\mathrm{fix}]$ from $\mathbb{E}_a[\mathrm{fix}]$, the analogous expectation under the uniform measure. Namely, as the action $\mathrm{GL}_N(K) \curvearrowright K^N$ admits two orbits (the zero vector and all nonzero vectors), the expected number of vectors in K^N fixed by a uniformly random element of $\mathrm{GL}_N(K)$ is $\mathbb{E}_a[\mathrm{fix}] = 2$, and we consider the difference $\mathbb{E}_w[\mathrm{fix}] - 2$. Theorems 1.1 and 1.2 imply that if w is a proper power, then $\mathbb{E}_w[\mathrm{fix}] - 2$ is of order $\Theta(1)$, and otherwise, it is of order $O\left(\frac{1}{q^N}\right)$. Next, we provide a more refined and accurate description of this difference in the nonpower case. To state our result and conjecture, we first define the notion of primitivity of elements in ideals. Recall that by Cohn and Lewin's result, every ideal $I \leq \mathcal{A}$ is a free $\mathcal{A}$-module and so admits a basis. All bases of I have the same cardinality, called the rank of I and denoted $\mathrm{rk}\, I$.

Definition 1.3. Let $I \leq \mathcal{A}$ be an ideal and let $f \in I$. We say that f is a *primitive* element of I if it is contained in some basis of I (considering I as a free right $\mathcal{A}$-module). Otherwise, f is *imprimitive* in I.

This is analogous to the notion of a primitive element in a free group: an element belonging to some basis of this group. Our next central result captures the $\frac{1}{q^N}$-term of the Laurent expansion of $\mathbb{E}_w[\mathrm{fix}]$.

Theorem 1.4. *Let $1 \neq w \in \boldsymbol{F}$ be a nonpower. Then the expected number of vectors in K^N fixed by a w-random element of $\mathrm{GL}_N(K)$ is*

$$\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{|\mathrm{Crit}_q^2(w)|}{q^N} + O\left(\frac{1}{q^{2N}}\right),$$

where $\mathrm{Crit}_q^2(w)$ is the set of ideals $I \leq \mathcal{A}$ of rank two which contain the element $w - 1$ as an imprimitive element.

As implied by the theorem, the set $\mathrm{Crit}_q^2(w)$ is indeed finite for every nonpower w. We prove this fact directly in Corollary 3.11. To illustrate, consider the commutator word $w = [a, b]$. As mentioned above,

$$\mathbb{E}_{[a,b]}[\mathrm{fix}] = 2 + \frac{(q-1)^2 q^N - (q-1)^3}{(q^N - 1)(q^N - q)} = 2 + \frac{(q-1)^2}{q^N} + O\left(\frac{1}{q^{2N}}\right).$$

[1] Some claim that the first correct proof of this result (stated formally below as Theorem 3.1) is due to Lewin [1969] — see [Hog-Angeloni 1990, Footnote 5].

[2] For example, it can be shown that the augmentation ideal $I_{\boldsymbol{F}} = \left\{\sum \alpha_w w \mid \sum \alpha_w = 0\right\} \subseteq \mathcal{A}$ is of rank $r = \mathrm{rk}\, \boldsymbol{F}$. For instance, when $\boldsymbol{F} = \boldsymbol{F}(a, b, c)$, $I_{\boldsymbol{F}} = (a-1)\mathcal{A} \oplus (b-1)\mathcal{A} \oplus (c-1)\mathcal{A}$.

In this case there are exactly $(q-1)^2$ distinct ideals of rank two containing $[a,b]-1$ as an imprimitive element: these are $(\delta a-1, \varepsilon b-1)$ with $\delta, \varepsilon \in K^*$. We conjecture a more general phenomenon, for which we make the following definition.

Definition 1.5. The *q-primitivity rank* of $w \in \boldsymbol{F}$, denoted $\pi_q(w)$, is the smallest rank of a *proper* ideal of $\mathcal{A}$ containing $w-1$ as an imprimitive element. Namely,

$$\pi_q(w) \stackrel{\text{def}}{=} \min\{\operatorname{rk} I \mid I \lneq \mathcal{A},\ I \ni w-1 \text{ and } w-1 \text{ is imprimitive in } I\}.$$

If this set is empty, we set $\pi_q(w)=\infty$. A *critical* ideal for w is a proper ideal of rank $\pi_q(w)$ containing $w-1$ as an imprimitive element. We denote by $\mathrm{Crit}_q(w)$ the set of critical ideals for w.

Corollary 3.17 shows that $\pi_q(w)$ takes values only in $\{0,1,\dots,r\}\cup\{\infty\}$, where r is the rank of $\boldsymbol{F}$. Note that $\pi_q(w)=0$ if and only if $w=1$: the only rank-0 ideal is (0), whose only basis is the empty set. In Section 4A below, we prove that $\pi_q(w)=1$ if and only if $w \in \boldsymbol{F}$ is a proper power (Corollary 4.6), and that in this case, if one writes $w=u^d$ with $d \geq 2$ and u a nonpower, the set of critical ideals of w is

$$\mathrm{Crit}_q(u^d) = \{(p(u)) : p \mid x^d-1 \in K[x],\ p \text{ monic and } p \neq 1, x^d-1\}.$$

For example, if $|K|=q=3$ and $w=u^4$, the critical ideals of w are in one-to-one correspondence with the six nontrivial monic divisors the polynomial $x^4-1 \in K[x]$. These rank-1 ideals are $(u-1)$, $(u+1)$, (u^2-1), (u^2+1), (u^3-u^2+u-1) and (u^3+u^2+u+1). The trivial monic divisors of x^4-1 correspond to the ideal $(1)=\mathcal{A}$ which is not proper, and to the ideal (u^4-1) in which $w-1$ is primitive. By Proposition 3.16, $\pi_q(w)=\infty$ if and only if w is a primitive element of $\boldsymbol{F}$.

The following conjecture thus generalizes Theorems 1.2 and 1.4.

Conjecture 1.6. *Let $w \in \boldsymbol{F}$ and let $\pi=\pi_q(w)$. Then the expected number of vectors in K^N fixed by a w-random element of $\mathrm{GL}_N(K)$ is*

$$\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{|\mathrm{Crit}_q(w)|}{q^{N\cdot(\pi-1)}} + O\left(\frac{1}{q^{N\cdot\pi}}\right). \tag{1-2}$$

Corollary 3.11 yields that $\mathrm{Crit}_q(w)$ is indeed finite. If $\pi := \pi_q(w) = 0$ (namely, if $w=1$), then $\mathrm{Crit}_q(w)=\{(0)\}$ and (1-2) is obvious. Theorem 1.2 proves (1-2) when $\pi=1$, and Theorem 1.4 proves it when $\pi=2$. As mentioned above, $\pi_q(w)=\infty$ if and only if w is primitive in $\boldsymbol{F}$, and in this case a w-random element of $\mathrm{GL}_N(K)$ distributes uniformly [Puder and Parzanchevski 2015, Observation 1.2], and so (1-2) holds. In particular, Conjecture 1.6 holds for the free group of rank 2 as the possible values of $\pi_q(w)$ are $\{0,1,2,\infty\}$ (Corollary 3.17). We conclude the following analog of a result about S_N [Puder 2014, Theorem 1.5].

Corollary 1.7. *Let $w \in \boldsymbol{F}_2$. Then w induces the uniform measure on $\mathrm{GL}_N(K)$ for all N if and only if w is primitive.*

Another important background for Conjecture 1.6 is an analogous result in the case of the symmetric group S_N. The primitivity rank of a word $w \in \boldsymbol{F}$, denoted $\pi(w)$ and introduced in [Puder 2014], is the

smallest rank of a subgroup of $\boldsymbol{F}$ containing w as an imprimitive element. Let $\mathrm{Crit}_{\boldsymbol{F}}(w)$ denote the set of subgroups of $\boldsymbol{F}$ of rank $\pi(w)$ which contain w as an imprimitive element. Then the S_N-analog of Conjecture 1.6 is [Puder and Parzanchevski 2015, Theorem 1.8]: the expected number of fixed points in a w-random permutation in S_N is

$$1 + \frac{|\mathrm{Crit}_{\boldsymbol{F}}(w)|}{N^{\pi(w)-1}} + O\left(\frac{1}{N^{\pi(w)}}\right).$$

Alongside its role in word measures on S_N, the original primitivity rank $\pi(w)$ seems to play a universal role in word measures on groups (see [Hanany and Puder 2023, Conjecture 1.13]), it has connections with stable commutator length (see Section 1.6 in the same article) and was recently found relevant to the study of one-relator groups (see, for example, [Louder and Wilton 2022]). Definition 1.5 seemingly introduces a family of related invariants of words — one for every prime power q. In fact, the same definition can be applied to arbitrary fields — see Section 7. However, it is possible that all these invariants coincide for a given word. We are able to show one inequality and conjecture a full equality.

Proposition 1.8. *For every word $w \in \boldsymbol{F}$ and every prime power q, $\pi_q(w) \le \pi(w)$.*

Conjecture 1.9. *For every word $w \in \boldsymbol{F}$ and every prime power q, $\pi_q(w) = \pi(w)$.*

Conjecture 1.9, along with Conjecture 1.6, are in line with a universal conjecture — [Hanany and Puder 2023, Conjecture 1.13] — about the role of the primitivity rank $\pi(w)$ in word measures on groups. For more background, see Section 1.6 in the same article.

As part of our study of word measures in $\mathrm{GL}_N(K)$ employing the free group algebras, we also prove some results about these algebras which may be of independent interest. For example, suppose that T is a subtree of the Cayley graph of $\boldsymbol{F}$ with respect to some basis. If $I \le \mathcal{A}$ is a finitely generated ideal with a generating set supported on T, then I admits a basis which is supported on T (Theorem 3.8). We also analyze the $\mathcal{A}$-module $\mathcal{A}/(w-1)$ obtained as the quotient of the right $\mathcal{A}$-module $\mathcal{A}$ by its submodule $(w-1)$. Theorem 5.4 proves an analog of Kaplansky's unit conjecture for these modules and shows that if w is a nonpower, then the only cyclic generators of $\mathcal{A}/(w-1)$ are the trivial ones. See Section 7 for a further discussion of this line of research.

1C. ***General stable class functions and characters.*** As mentioned above, some of the results concerning the function fix and its expectation under word measures are only an illustrative special case of more general results. The variety of functions we consider are those relating to *stable* representations of the family $\mathrm{GL}_\bullet(K)$ (see [Putman and Sam 2017; Gan and Watterlond 2018]). Below we present the generalizations of Theorems 1.1 and 1.2 and of Conjecture 1.6.

First, we must remark on the unconventional definition we make in this paper. Formal words in group theory are usually read from left to right: this is why one usually considers *right* Cayley graphs. As a consequence, we consider here the slightly nonstandard *right* action of $\mathrm{GL}_N(K)$ on $V_N \overset{\mathrm{def}}{=} K^N$, namely, we consider V_N as row vectors, and the action of $g \in \mathrm{GL}_N(K)$ on $v \in V_N$ is given by $(v, g) \mapsto vg$. Thus,

the action of $w(g_1, \dots, g_r)$ on a vector $v \in V_N$ can be thought of as the composition of the action, letter by letter, from left to right — the natural direction in which the word is read.

Rather than considering only the number of vectors fixed by g, we consider more generally the number of subspaces of V of a fixed dimension which are invariant under g and on which g acts in a prescribed way. This is formalized as follows:

Definition 1.10. Let $m \in \mathbb{Z}_{\geq 1}$ and $\mathcal{B} \in \mathrm{GL}_m(K)$. We define a map $\tilde{\mathcal{B}} : \mathrm{GL}_N(K) \to \mathbb{Z}_{\geq 0}$ (valid for arbitrary N) as follows. For $g \in \mathrm{GL}_N(K)$ we let $\tilde{\mathcal{B}}(g)$ be the number of m-tuples of vectors $v_1, \dots, v_m \in V_N = K^N$ on which the (right) action of g can be described by a multiplication from the left by the matrix $\mathcal{B}$. Namely,

$$\tilde{\mathcal{B}}(g) = \#\{M \in M_{m\times N}(K) \mid Mg = \mathcal{B}M\}.$$

For example, if $\mathcal{B} = (1) \in \mathrm{GL}_1(K)$, then $\tilde{\mathcal{B}} = \mathrm{fix}$. For $\mathcal{B} = (\lambda) \in \mathrm{GL}_1(K)$, the function $\tilde{\mathcal{B}}$ gives the size of the eigenspace $V_\lambda \leq V_N$ of an element. If $\mathcal{B} = I_m \in \mathrm{GL}_m(K)$, then $\tilde{\mathcal{B}}(g) = \mathrm{fix}(g)^m$, and if

$$\mathcal{B} = \begin{pmatrix} & 1 & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \\ 1 & & & & \end{pmatrix} \in \mathrm{GL}_m(K),$$

then $\tilde{\mathcal{B}}(g) = \mathrm{fix}(g^m)$. The following two theorems are the generalization of Theorems 1.1 and 1.2:

Theorem 1.11. *Suppose that $w \in \boldsymbol{F}$, $m \in \mathbb{Z}_{\geq 1}$ and $\mathcal{B} \in \mathrm{GL}_m(K)$. Then for every large enough N, the expectation $\mathbb{E}_w[\tilde{\mathcal{B}}]$ is given by a rational function in q^N.*

Theorem 1.12. *Let $1 \neq w = u^d$ with $d \geq 1$ and u a nonpower. For every $m \in \mathbb{Z}_{\geq 1}$ and $\mathcal{B} \in \mathrm{GL}_m(K)$, the limit $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$ exists and depends only on d and not on u.*

In the special case of $\tilde{\mathcal{B}} = I_m \in \mathrm{GL}_m(K)$, Theorem 1.11 appeared in [Ernst-West 2019]. The same special case of Theorem 1.12 first appeared in the same thesis, and then, independently, in [Eberhard and Jezernik 2022, Section 8].

In particular, Theorem 1.12 captures all moments of the number of fixed vectors under the w-measure. So if $w = u^d$, all these moments converge, as $N \to \infty$, to the same limits as for $w = a^d$, namely as for a d-th power of a uniformly random element of $\mathrm{GL}_N(K)$. Denote the number of fixed vectors in K^N of a w-random element of $\mathrm{GL}_N(K)$ by $\mathrm{fix}_{w,N}$. When $w = a$, a limit distribution as $N \to \infty$ is known to exist [Fulman and Stanton 2016, Theorem 2.1].[3] Although this limit distribution is not determined by its moments, we do prove the following in the Appendix:

Theorem 1.13. *Let $1 \neq w \in \boldsymbol{F}$ be a nonpower. Then the random variables $\mathrm{fix}_{w,N}$ have a limit distribution, and this limit distribution is identical to the one of $\mathrm{fix}_{a,N}$ described in* [Fulman and Stanton 2016, Theorem 2.1].

[3] This was originally proved by Rudvalis and Shinoda — see [Fulman and Stanton 2016].

Theorem 1.13 is analogous to the $L = d = 1$ case of Nica's main Theorem 1.1 [1994], which revolves around the limit distribution of the number of fixed point in w-random permutations. We suspect that Theorem 1.13 can be generalized to a full analog of Nica's result (and see Section 7).

Remark 1.14. One can further generalize Theorem 1.11 to more than one word. For example, for any tuple of words $w_1, \ldots, w_\ell \in \boldsymbol{F}$, consider an ℓ-tuple of random elements

$$\overline{w_1} = w_1(g_1, \ldots, g_r), \ldots, \overline{w_\ell} = w_\ell(g_1, \ldots, g_r) \in \mathrm{GL}_N(K),$$

where $g_1, \ldots, g_r$ are independent, uniformly random elements of $\mathrm{GL}_N(K)$, and consider expressions like $\mathbb{E}[\mathrm{fix}(\overline{w_1}) \cdot \mathrm{fix}(\overline{w_2}) \cdots \mathrm{fix}(\overline{w_\ell})]$. The same argument given in the proof of Theorem 1.11 shows that this expectation is given by a rational expression in q^N. Also, Corollary 1.3 in [Ernst-West 2019] shows that the difference $\mathbb{E}[\mathrm{fix}(\overline{w_1}) \cdots \mathrm{fix}(\overline{w_\ell})] - \mathbb{E}_{w_1}[\mathrm{fix}] \cdots \mathbb{E}_{w_\ell}[\mathrm{fix}] = O\left(\frac{1}{q^N}\right)$ if and only if no pair of words is conjugated into the same cyclic subgroup of $\boldsymbol{F}$.

We further introduce a generalization of Conjecture 1.6. Consider

$$\mathcal{R} \stackrel{\mathrm{def}}{=} \mathbb{C}\left[\{\tilde{\mathcal{B}} \mid \mathcal{B} \in \mathrm{GL}_m(K),\ m \in \mathbb{Z}_{\geq 0}\}\right],$$

the $\mathbb{C}$-algebra generated by all functions $\tilde{\mathcal{B}}$ from Definition 1.10. Every element of $\mathcal{R}$ is a (class) function defined on $\mathrm{GL}_N(K)$ for every N. Rather than formal polynomials in the $\tilde{\mathcal{B}}$'s, the elements of $\mathcal{R}$ are functions on $\mathrm{GL}_\bullet(K)$, so two elements giving the same function on $\mathrm{GL}_N(K)$ for every N are identified. For example, every conjugate of $\mathcal{B}$ gives rise to the same function as $\mathcal{B}$. In fact, this is the only case where two elements give the same function: $\tilde{\mathcal{B}}_1 = \tilde{\mathcal{B}}_2$ if and only if $\mathcal{B}_1$ and $\mathcal{B}_2$ belong to $\mathrm{GL}_m(K)$ for the same m and are conjugates — see [Ernst-West et al. 2024, Corollary 3.1]. If we also include the constant function 1, thought of as $\tilde{\mathcal{B}}$ where $\mathcal{B} = e \in \mathrm{GL}_0(K) \stackrel{\mathrm{def}}{=} \{e\}$, then $\mathcal{R}$ is the $\mathbb{C}$-span of the $\tilde{\mathcal{B}}$'s: indeed, if $\mathcal{B}_1 \in \mathrm{GL}_{m_1}(K)$ and $\mathcal{B}_2 \in \mathrm{GL}_{m_2}(K)$, then $\tilde{\mathcal{B}}_1 \cdot \tilde{\mathcal{B}}_2 = \widetilde{\mathcal{B}_1 \oplus \mathcal{B}_2}$ where $\mathcal{B}_1 \oplus \mathcal{B}_2 \in \mathrm{GL}_{m_1+m_2}(K)$ is the suitable block-diagonal matrix. In the same article, it is shown that $\mathcal{R}$ is, in fact, a graded algebra and admits a linear basis consisting of $\{\tilde{\mathcal{B}}\}$, where $\mathcal{B}$ goes over exactly one representative from every conjugacy class in all $\mathrm{GL}_m(K)$ $(m \geq 0)$.

Some of the functions in $\mathcal{R}$ coincide, for large enough N, with irreducible characters of $\mathrm{GL}_N(K)$. For example, for $N \geq 2$, the action of $\mathrm{GL}_N(K)$ on the projective space $\mathbb{P}^{N-1}(K)$ decomposes to the trivial representation and an irreducible representation whose character we denote $\chi^{\mathbb{P}}$. Then for every $N \geq 2$, the character $\chi^{\mathbb{P}}$ is equal to an element in $\mathcal{R}$:

$$\chi^{\mathbb{P}} = \frac{1}{q-1} \sum_{\lambda \in K^*} (\tilde{\lambda} - 1) - 1$$

(here $\tilde{\lambda}$ is the function corresponding to $\lambda \in \mathrm{GL}_1(K)$). In [Ernst-West et al. 2024] it is shown that the set of families of irreducible characters $\{\chi_N \in \mathrm{GL}_N(K)\}_{N \geq N_0}$ which coincide with elements of $\mathcal{R}$ is precisely the set of *stable* irreducible representations of $\mathrm{GL}_\bullet(K)$ as in [Gan and Watterlond 2018]. Our generalization of Conjecture 1.6 deals with these families of irreducible characters.

Conjecture 1.15. *Let χ be a stable character of $\mathrm{GL}_\bullet(K)$, namely, an element of $\mathcal{R}$ which coincides, for every large enough N, with some irreducible character of $\mathrm{GL}_N(K)$. Then*

$$\mathbb{E}_w[\chi] = O((\dim \chi)^{1-\pi_q(w)}).$$

By Theorem 1.11 (with $w = 1$), $\dim \chi = \chi(1)$ is a rational function in q^N. Conjecture 1.15, together with a positive answer to Conjecture 1.9, constitute a special case of the more general, albeit not as precise, [Hanany and Puder 2023, Conjecture 1.13]. (See also [Ernst-West et al. 2023, Conjecture A.4] for a slightly more ambitious version of Conjecture 1.15.)

For $N \geq 2$, the decomposition of the function fix to irreducible characters is

$$\mathrm{fix} = 2 \cdot \mathrm{triv} + \chi^{\mathbb{P}} + \xi_1 + \cdots + \xi_{q-2},$$

where $\xi_1, \ldots, \xi_{q-2}$ are distinct irreducible characters, each of dimension $\frac{q^N-1}{q-1}$, all belonging to $\mathcal{R}$. Thus, they all fall into the framework of Conjecture 1.15, and we get that this conjecture implies, in particular, that $\mathbb{E}_w[\mathrm{fix}] = 2 + O((q^N)^{1-\pi_q(w)})$. In particular, Conjecture 1.15 generalizes (a slightly weaker version of) Conjecture 1.6. Some background for Conjecture 1.15 can be found in [Hanany and Puder 2023, Section 1].

1D. *Reader's guide.*

Notation. The free group $\boldsymbol{F}$ has rank r and a fixed basis $B = \{b_1, \ldots, b_r\}$. Recall that all ideals in this paper are one-sided right ideals unless stated otherwise, and we write $I \leq \mathcal{A}$ to mean that I is an ideal of the free group algebra $\mathcal{A} = K[\boldsymbol{F}]$. More generally, we write $M \leq \mathcal{A}^m$ if M is a submodule of the free right $\mathcal{A}$-module $\mathcal{A}^m$. For any set $S \subseteq \mathcal{A}^m$, we denote by (S) the submodule generated by S, and if $S = \{s_1, \ldots, s_t\}$ we may also write $(s_1, \ldots, s_t)$.

We denote by $E = \{e_1, \ldots, e_m\}$ a basis for the free $\mathcal{A}$-module $\mathcal{A}^m$. The elements of the form ez with $e \in E$ and $z \in \boldsymbol{F}$ are called *monomials*. For a subset Q of the monomials, we write $M \leq_Q \mathcal{A}^m$ to mean that M has a generating set in $\mathcal{A}^m$ such that each of its elements is supported on Q. Usually, for ideals inside $\mathcal{A}$, we consider subsets of $\boldsymbol{F}$ corresponding to the vertices in some subtree T of the (right) Cayley graph $\mathcal{C} \stackrel{\text{def}}{=} \mathrm{Cay}(\boldsymbol{F}, B)$ of $\boldsymbol{F}$ with respect to the basis B. In this case, instead of $I \leq_{\mathrm{vert}(T)} \mathcal{A}$ (here, of course, $\mathrm{vert}(T)$ denotes the set of vertices of T), we simply write $I \leq_T \mathcal{A}$. More generally, for submodules of $\mathcal{A}^m$, we usually consider m disjoint copies $\mathcal{C}_1, \ldots, \mathcal{C}_m$ of $\mathrm{Cay}(\boldsymbol{F}, B)$, with origins $e_1, \ldots, e_m$, respectively, and consider a collection of (possibly empty) subtrees $\mathbb{T} = T_1 \cup \cdots \cup T_m$, with $T_i \subset \mathcal{C}_i$. We write $M \leq_{\mathbb{T}} \mathcal{A}^m$ to mean that M is generated by elements supported on the vertices of $\mathbb{T}$.

For a submodule $M \leq \mathcal{A}^m$ and a set S of monomials in $\mathcal{A}^m$, we let $M|_S$ denote the set of elements of M which are supported on S. This is a vector space over K.

Paper organization. After a very brief survey of related works in Section 1E, Section 2 proves that $\mathbb{E}_w[\mathrm{fix}]$, and likewise $\mathbb{E}_w[\tilde{\mathcal{B}}]$, are given by rational functions in q^N (Theorems 1.1 and 1.11, respectively). In Section 3 we study the free group algebra and its ideals, show how the computation of $\mathbb{E}_w[\mathrm{fix}]$ is related to "exploration" processes in the Cayley graph of $\boldsymbol{F}$, and prove some basic properties of the q-primitivity rank including Proposition 1.8. We then study $\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}]$ and $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$ and prove Theorems 1.2 and

1.12 in Section 4. Section 5 studies the right $\mathcal{A}$-module $\mathcal{A}/(w-1)$ and specifies its cyclic generators, and also gives a criterion to detect when $w-1$ is primitive in a given rank-2 ideal in $\mathcal{A}$. Section 6 deals with the coefficient of $\frac{1}{q^N}$ in the Laurent expansion of $\mathbb{E}_w[\mathrm{fix}]$ and proves Theorem 1.4. Section 7 gathers the many open questions that are raised by this work. Finally, the Appendix contains the proof of Theorem 1.13.

1E. ***Related works.*** As mentioned above, the two phenomena described in Theorems 1.11 and 1.12 are found in other families of groups. The fact that the expectation under word measures of "natural" class functions over certain families of groups are given by rational functions was first established for the symmetric group [Nica 1994; Linial and Puder 2010]. It was later established for the classical groups $U(N)$ [Magee and Puder 2019] and $O(N)$ and $\mathrm{Sp}(N)$ [Magee and Puder 2024] based on Weingarten calculus (see, for instance, [Collins and Śniady 2006]), and also in the wreath product $G \wr S_N$ for an arbitrary finite group G [Magee and Puder 2021; Shomroni 2023a]. A related phenomenon appears when free groups are replaced by surface groups (fundamental groups of compact closed surfaces). Indeed, there is a natural definition of a measure induced by an element of a surface group on finite groups and certain compact groups, and the expected value of certain characters of the symmetric group $\mathrm{Sym}(N)$ under such measures can be approximated to any degree by a rational function [Magee and Puder 2023]. A similar result holds for measures induced by elements of surface groups on $\mathrm{SU}(N)$ [Magee 2022].

The phenomenon described in Theorems 1.2 and 1.12, that if $w = u^d$ then the limit expectation of natural class functions in the family under the w-measure depends only on d and not on u, is also found in many of the above mentioned cases. It is true in $\mathrm{Sym}(N)$ [Nica 1994; Linial and Puder 2010], in $U(N)$ [Mingo et al. 2007; Rădulescu 2006], as well as in $O(N)$ and in $\mathrm{Sp}(N)$ [Magee and Puder 2024]. It also holds in the characters analyzed in [Magee and Puder 2023] for measures on $\mathrm{Sym}(N)$ induced by elements of surface groups [ibid., Theorem 1.2].

Finally, there are analogs to Theorem 1.4 and Conjectures 1.6 and 1.15, which give an interpretation to the order of $\mathbb{E}_w[f] - \mathbb{E}_x[f]$, an interpretation which lies in invariants of w as an element of the abstract free group $\boldsymbol{F}$. We mentioned previously that there are very similar results in the case of $\mathrm{Sym}(N)$ [Puder and Parzanchevski 2015; Hanany and Puder 2023]. There are other invariants of w explaining the leading order (and sometimes much more than the leading order) in the expected values of class functions in $U(N)$, $O(N)$, $\mathrm{Sp}(N)$ and $G \wr S_N$ [Magee and Puder 2019; 2021; 2024; Brodsky 2024; Shomroni 2023a; 2023b]. A more detailed summary may be found in [Hanany and Puder 2023, Section 1.3].

2. Rational expressions

We prove Theorems 1.1 and 1.11, which show that the expectations under word measures of the class functions we consider on $\mathrm{GL}_N(K)$ are given by rational functions in q^N. The proof uses only linear algebra and can be written in completely elementary terms. While we start with this approach, we then "translate" the proof to the language of ideals and modules of the free group algebra $\mathcal{A} = K[\boldsymbol{F}]$. Given our additional results and conjectures, the latter language is much more fruitful.

2A. ***The function* fix *and Theorem 1.1.***

The elementary approach. We first illustrate the proof in the somewhat simpler special case considered in Theorem 1.1: the function fix. Let $V_N = K^N$ be the vector space of row vectors of length N. Given $w \in \boldsymbol{F}$, one needs to count all $g_1, \dots, g_r \in \mathrm{GL}_N(K)$ and $v \in V_N$ such that $v.w(g_1, \dots, g_r) = v$. We consider the entire trajectory of v when the letters of w are applied one by one. Namely, assume that w is written in the basis $B = \{b_1, \dots, b_r\}$ of $\boldsymbol{F}$ as $w = b_{i_1}^{\varepsilon_1} \dots b_{i_\ell}^{\varepsilon_\ell}$ (where $i_j \in \{1, \dots, r\}$ and $\varepsilon_j \in \{\pm 1\}$). We consider the vectors

$$v^0 \stackrel{\text{def}}{=} v, \quad v^1 \stackrel{\text{def}}{=} v^0.g_{i_1}^{\varepsilon_1}, \quad v^2 \stackrel{\text{def}}{=} v^1.g_{i_2}^{\varepsilon_2}, \quad \dots, \quad v^{\ell-1} \stackrel{\text{def}}{=} v^{\ell-2}.g_{i_{\ell-1}}^{\varepsilon_{\ell-1}}, \quad v^\ell \stackrel{\text{def}}{=} v^{\ell-1}.g_{i_\ell}^{\varepsilon_\ell} = v^0. \tag{2-1}$$

We denote this trajectory by $\bar{v} = (v^0, \dots, v^\ell)$. Given that the entire trajectory is determined by $g_1, \dots, g_r$ and $v = v^0$, we do not change our goal by counting $(g_1, \dots, g_r; \bar{v})$ satisfying the equations in (2-1) instead of $(g_1, \dots, g_r; v)$ satisfying $v.w(g_1, \dots, g_r) = v$.

The basic idea behind our counting is grouping together solutions $(g_1, \dots, g_r; \bar{v})$ according to the linear relations over K which $v^0, \dots, v^\ell$ satisfy. There are finitely many options here (trivially, at most the number of linear subspaces of $K^{\ell+1}$), and, as we show below, for each subspace of $K^{\ell+1}$ the number of solutions $(g_1, \dots, g_r; \bar{v})$ corresponding to it is either identically zero for every N, or its contribution to $\mathbb{E}_w[\mathrm{fix}]$ is given by a rational function in q^N for every large enough N.

Denote by $[1, w]$ the subtree of $\mathcal{C} = \mathrm{Cay}(\boldsymbol{F}, B)$ corresponding to the path from the origin to the vertex w. For every $b \in B$, denote by $D_b(w)$ the vertices of $[1, w]$ with an outgoing b-edge (within $[1, w]$), and denote by $e_b(w)$ the number of b-edges in $[1, w]$, so $e_b(w) = |D_b(w)|$. Now consider a subspace $\Delta \le K^{\ell+1}$ thought of as a set of equations on the vectors $v^0, \dots, v^\ell$, or, equivalently, on the vertices of $[1, w]$. Below we denote these vertices by the corresponding prefix of w in $\boldsymbol{F}$, and write elements of $K^{[1,w]} \stackrel{\text{def}}{=} K^{\mathrm{vert}([1,w])} \cong K^{\ell+1}$ as linear combinations of these vertices. We have

$$\begin{aligned}
\mathbb{E}_w[\mathrm{fix}] &= \frac{\#\{g_1, \dots, g_r \in \mathrm{GL}_N(K),\, v \in V_N \mid v.w(g_1, \dots, g_r) = v\}}{|\mathrm{GL}_N(K)|^r} \\
&= \frac{\#\{g_1, \dots, g_r \in \mathrm{GL}_N(K),\, \bar{v} \in V_N^{\ell+1} \mid \bar{v} \text{ and } g_1, \dots, g_r \text{ satisfy (2-1)}\}}{|\mathrm{GL}_N(K)|^r} \\
&= \sum_{\Delta \le K^{[1,w]}} \frac{\#\{g_1, \dots, g_r \in \mathrm{GL}_N(K),\, \bar{v} \in V_N^{\ell+1} \mid \bar{v} \text{ satisfies precisely } \Delta,\, \bar{v}, g_1, \dots, g_r \text{ satisfy (2-1)}\}}{|\mathrm{GL}_N(K)|^r}
\end{aligned} \tag{2-2}$$

If there are solutions $(g_1, \dots, g_r; \bar{v})$ which satisfy precisely Δ, then the following two conditions hold:

C1: $w - 1 \in \Delta$ (here $w - 1$ is the equation $w - 1 = 0$, or, equivalently, $v^\ell - v^0 = 0$).

C2: Δ is "closed under multiplication by $b^{\pm 1}$". Namely, for every $b \in B$ and every equation $\delta = \sum_{z \in D_b(w)} \lambda_z z$ ($\lambda_z \in K$) supported on $D_b(w)$, denote by $\delta b \stackrel{\text{def}}{=} \sum_{z \in D_b(w)} \lambda_z z b$ the corresponding equation on the vertices on the termini of the corresponding b-edges. Then

$$\delta \in \Delta \iff \delta b \in \Delta.$$

Conversely, if Δ satisfies conditions **C1** and **C2**, then for every large enough N there exist solutions $(g_1, \dots, g_r; \bar{v})$ satisfying precisely Δ, and the contribution of Δ in (2-2) is given by a rational function

in q^N. Indeed, denote by $\dim(\Delta)$ the dimension of the subspace Δ, and by $\dim_b(\Delta)$ the dimension of the subspace of Δ consisting of equations supported on $D_b(w)$. First, we choose a trajectory $\bar{v} \in V_N^{\ell+1}$ satisfying precisely Δ. The number of choices for such $\bar{v}$ is precisely $\mathfrak{indep}_{\ell+1-\dim(\Delta)}(V_N)$, where

$$\mathfrak{indep}_h(V_N) \stackrel{\text{def}}{=} (q^N-1)(q^N-q)\cdots(q^N-q^{h-1})$$

is the number of h-tuples of independent vectors in V_N.[4]

Second, given a trajectory $\bar{v}$ satisfying precisely Δ, we choose the tuple $g_1,\dots,g_r \in \mathrm{GL}_N(K)$ so that $\bar{v}, g_1,\dots,g_r$ satisfy (2-1). We choose g_i separately for every $i=1,\dots,r$. Let $b=b_i$. The vectors of $\bar{v}$ at the starting points of b-edges in $[1,w]$, namely, in $D_b(w)$, span a subspace of V of dimension $e_b(w)-\dim_b(\Delta)$. (Such a trajectory may exist only if $e_b(w)-\dim_b(\Delta)\le N$). In this case, the element g_i should map a subspace of dimension $e_b(w)-\dim_b(\Delta)$ in a prescribed way, and condition **C2** guarantees this prescribed way is valid and can be realized by a linear transformation. The number of elements in $\mathrm{GL}_N(K)$ satisfying this constraint is

$$(q^N-q^{e_b(w)-\dim_b(\Delta)})(q^N-q^{e_b(w)-\dim_b(\Delta)+1})\cdots(q^N-q^{N-1}).$$

If $g_1,\dots,g_r$ satisfy these constraints and as **C1** holds, $\bar{v}$ and $g_1,\dots,g_r$ satisfy (2-1). Overall, if $N \ge e_b(w)-\dim_b(\Delta)$ for every $b\in B$, the term corresponding to Δ in (2-2) is

$$\mathfrak{indep}_{\ell+1-\dim(\Delta)}(V_N)\cdot\prod_{b\in B}\frac{(q^N-q^{e_b(w)-\dim_b(\Delta)})(q^N-q^{e_b(w)-\dim_b(\Delta)+1})\cdots(q^N-q^{N-1})}{(q^N-1)(q^N-q)\cdots(q^N-q^{N-1})}$$
$$=\frac{\mathfrak{indep}_{\ell+1-\dim(\Delta)}(V_N)}{\prod_{b\in B}\mathfrak{indep}_{e_b(w)-\dim_b(\Delta)}(V_N)},$$

which is rational in q^N. Overall, we obtain

$$\mathbb{E}_w[\mathrm{fix}] = \sum_{\Delta\le K^{[1,w]}:\,\Delta \text{ satisfies } \mathbf{C1},\mathbf{C2}} \frac{\mathfrak{indep}_{\ell+1-\dim(\Delta)}(V_N)}{\prod_{b\in B}\mathfrak{indep}_{e_b(w)-\dim_b(\Delta)}(V_N)}, \tag{2-3}$$

which completes the proof of Theorem 1.1.

The free-group-algebra approach. The key observation that leads to the free-group-algebra approach is that condition **C2** above is a feature of (as always, right) ideals of the free group algebra $\mathcal{A}=K[\boldsymbol{F}]$: a right ideal $I\le\mathcal{A}$ is a K-linear subspace of $\mathcal{A}$ satisfying **C2** on the entire Cayley graph $\mathcal{C}$ (rather than on $[1,w]$ alone). To make this formal, let us recall some notation. For a subtree T of the Cayley graph $\mathcal{C}=\mathrm{Cay}(\boldsymbol{F},B)$, denote by $D_b(T)$ the set of vertices in the subtree $T\subset\mathcal{C}$ with an outgoing b-edge (inside T), and by $e_b(T)=|D_b(T)|$ the number of such edges. For any ideal $I\le\mathcal{A}$, its restriction to T, denoted

$$I|_T \stackrel{\text{def}}{=} I\cap K^{\mathrm{vert}(T)},$$

[4]We could not find a conventional notation for the quantity $\mathfrak{indep}_h(V_N)$, but it is closely related to existing common notation. For example, $\mathfrak{indep}_h(v)=q^{Nh}\cdot(q^{-N};q)_h$, where $(t;q)_h \stackrel{\text{def}}{=} (1-t)(1-tq)\cdots(1-tq^{h-1})$ is the q-shifted factorial.

is a linear subspace of $K^{\mathrm{vert}(T)}$. We say that a K-linear subspace $\Delta \le K^{\mathrm{vert}(T)}$ satisfies $\mathbf{C2}(T)$ if for every $\delta \in K^{\mathrm{vert}(T)}$ supported on $D_b(T)$, we have $\delta \in \Delta$ if and only if $\delta.b \in \Delta$.

Lemma 2.1. *Assume that $\Delta \le K^{\mathrm{vert}(T)}$ is a K-linear subspace satisfying $\mathbf{C2}(T)$. Then $(\Delta) \le \mathcal{A}$, the ideal generated by Δ, does not introduce any new elements supported on T, namely*

$$(\Delta)|_T = \Delta. \tag{2-4}$$

Proof. It is clear that $(\Delta)|_T \supseteq \Delta$, so it is enough to show the converse inclusion. We may assume that T is finite: Every element of $\mathcal{A}$ has finite support, and every element of (Δ) is generated by finitely many elements of Δ. So if T is not finite and (2-4) fails, replace T with the finite subtree $S \subseteq T$ which is the convex hull of the support of an element in $(\Delta)|_T \setminus \Delta$ and its finitely many generators in Δ and replace Δ with $\Delta|_S$.

As in the proof of Theorem 1.1 above, for large enough N, there are $g_1, \dots, g_r \in \mathrm{GL}_N(K)$ and $\bar{v} = \{v_z \in V_N\}_{z \in \mathrm{vert}(T)}$ such that for every b-edge $z_1 \xrightarrow{b} z_2$ in T, we have $v_{z_1}.g_b = v_{z_2}$, and such that the equations over K satisfied by the vectors $\bar{v}$ are *precisely* the elements of Δ. The tuple $g_1, \dots, g_r$ defines a group homomorphism $\boldsymbol{F} \to \mathrm{GL}_N(K)$ by $b_i \mapsto g_i$. This group homomorphism defines, in turn, a homomorphism of K-algebras $\mathcal{A} \to \mathrm{End}(V_N)$. Equivalently, such a homomorphism of K-algebras defines a structure of an $\mathcal{A}$-module on V_N. Pick an arbitrary $z_0 \in \mathrm{vert}(T) \subseteq \boldsymbol{F}$. Now $\mathcal{A}$ is itself an $\mathcal{A}$-module and, moreover, it is a free $\mathcal{A}$-module with basis $\{z_0\}$. There is a unique $\mathcal{A}$-module homomorphism $\phi : \mathcal{A} \to V_N$ such that $\phi(z_0) = v_{z_0}$. Since ϕ is an $\mathcal{A}$-module homomorphism and T is connected, the choice of the g_i's guarantees that $\phi(z) = v_z$ for every $z \in \mathrm{vert}(T) \subseteq \boldsymbol{F}$.

Finally, $\ker\phi \le \mathcal{A}$ is a submodule, or an ideal, and the equations it satisfies on $\mathrm{vert}(T)$ are precisely those satisfied by $\bar{v}$, namely, Δ. Thus

$$(\Delta)|_T \le [\ker\phi]|_T = \Delta. \qquad \square$$

Returning to the case $T = [1, w]$, recall that we write $I \le_{[1,w]} \mathcal{A}$ if I is an ideal of $\mathcal{A}$ with generating set supported on $[1, w]$. Lemma 2.1 yields that there is a one-to-one correspondence

$$\{\Delta \le K^{[1,w]} \text{ satisfying } \mathbf{C1} \text{ and } \mathbf{C2}\} \iff \{I \le_{[1,w]} \mathcal{A} \mid w - 1 \in I\}.$$

For an ideal $I \le \mathcal{A}$ and every finite subtree T of $\mathcal{C}$, define

$$d^T(I) \stackrel{\mathrm{def}}{=} \dim_K(I|_T).$$

Similarly, for every basis element $b \in B$, denote by $D_b(T)$ the set of vertices in the subtree $T \subset \mathcal{C}$ with an outgoing b-edge (inside T), and let

$$d_b^T(I) \stackrel{\mathrm{def}}{=} \dim_K(I|_{D_b(T)}).$$

With this notation, (2-3) is equivalent to

$$\mathbb{E}_w[\mathrm{fix}] = \sum_{I \le_{[1,w]} \mathcal{A} : I \ni w-1} \frac{\mathfrak{indep}_{|w|+1-d^{[1,w]}(I)}(V_N)}{\prod_{b \in B} \mathfrak{indep}_{e_b(w)-d_b^{[1,w]}(I)}(V_N)}. \tag{2-5}$$

The advantage of translating (2-3) to the language of ideals as in (2-5) will soon be apparent. For example, Corollary 3.9 below shows that the summand in (2-5) corresponding to $I \leq_{[1,w]} \mathcal{A}$ is of order $(q^N)^{1-\mathrm{rk}\, I}$.

2B. ***The general case: Theorem 1.11.*** Fix $w \in \boldsymbol{F}$, $m \in \mathbb{Z}_{\geq 1}$ and $\mathcal{B} \in \mathrm{GL}_m(K)$. Our goal is to prove that for every large enough N, the expectation $\mathbb{E}_w[\tilde{\mathcal{B}}]$ is a rational function in q^N. Now we need to count tuples $v_1, \ldots, v_m \in V_N$ and $g_1, \ldots, g_r \in \mathrm{GL}_N(K)$ such that, defining $u_i \stackrel{\mathrm{def}}{=} v_i.w(g_1, \ldots, g_r)$, we have

$$\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \mathcal{B} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}. \tag{2-6}$$

As above, we consider the entire trajectories of $v_1, \ldots, v_m$ through the letters of w, namely,

$$\begin{array}{llll} v_1^0 = v_1, & v_1^1 = v_1.g_{i_1}^{\varepsilon_1}, & \ldots, & v_1^\ell = v_1.w(g_1, \ldots, g_r), \\ & \vdots & & \\ v_m^0 = v_m, & v_m^1 = v_m.g_{i_1}^{\varepsilon_1}, & \ldots, & v_m^\ell = v_m.w(g_1, \ldots, g_r), \end{array}$$

which we denote by $\bar{v}$. Again we group the solutions $(g_1, \ldots, g_r; \bar{v})$ according to the equations over K satisfied by $\bar{v}$. This time, the equations are not given by ideals in $\mathcal{A}$, but rather by submodules of the right free $\mathcal{A}$-module $\mathcal{A}^m$. Formally, let $E = \{e_1, \ldots, e_m\}$ be a basis of the free module $\mathcal{A}^m$. Every element of $\mathcal{A}^m$ is a finite linear combination, with coefficients from K, of monomials ez with $e \in E$ and $z \in \boldsymbol{F}$. These monomials are identified with the vertices of m disjoint copies $\mathcal{C}_1, \ldots, \mathcal{C}_m$ of $\mathrm{Cay}(\boldsymbol{F}, B)$, with origins $e_1, \ldots, e_m$, respectively.

Let $\mathbb{W}$ denote the union of the paths $[1, w]$ in $\mathcal{C}_1, \ldots, \mathcal{C}_m$, so $\mathbb{W} = \bigcup_{e \in E} [e, ew]$. Recall that $M \leq_{\mathbb{W}} \mathcal{A}^m$ means that M is a submodule of $\mathcal{A}^m$ with a generating set supported on $\mathbb{W}$. If the equations satisfied by the trajectory $\bar{v}$ are precisely $M|_{\mathbb{W}}$, then M must, in particular, contain the elements dictated by (2-6), which we denote by $\mathrm{EQ}_{\mathcal{B},w} \subseteq \mathcal{A}^m$. For example, if $\mathcal{B} = \left(\begin{smallmatrix} 2 & 1 \\ 7 & 3 \end{smallmatrix}\right) \in \mathrm{GL}_2(K)$, then $\mathrm{EQ}_{\mathcal{B},w}$ equals $\{e_1 w - 2e_1 - e_2, e_2 w - 7e_1 - 3e_2\}$.

Generalizing the notation from above, if $\mathbb{T} = T_1 \cup \cdots \cup T_m$ is a union of (possibly empty) subtrees $T_i \subseteq \mathcal{C}_i$, and $M \leq \mathcal{A}^m$, define

$$\begin{aligned} d^{\mathbb{T}}(M) &\stackrel{\mathrm{def}}{=} \dim_K(M|_{\mathbb{T}}), \\ d_b^{\mathbb{T}}(M) &\stackrel{\mathrm{def}}{=} \dim_K(M|_{D_b(\mathbb{T})}), \quad b \in B, \\ e_b(\mathbb{T}) &\stackrel{\mathrm{def}}{=} |D_b(\mathbb{T})|. \end{aligned}$$

So $|D_b(\mathbb{W})| = m \cdot e_b(w)$. The same argument as above shows that for every $N \geq \max_{b \in B} e_b(\mathbb{W})$,

$$\mathbb{E}_w[\tilde{\mathcal{B}}] = \sum_{M \leq_{\mathbb{W}} \mathcal{A}^m : M \supseteq \mathrm{EQ}_{\mathcal{B},w}} \frac{\mathfrak{indep}_{m(|w|+1) - d^{\mathbb{W}}(M)}(V_N)}{\prod_{b \in B} \mathfrak{indep}_{e_b(\mathbb{W}) - d_b^{\mathbb{W}}(M)}(V_N)}. \tag{2-7}$$

As there are finitely many submodules $M \leq_{\mathbb{W}} \mathcal{A}^m$, the expression (2-7) is rational in q^N. This completes the proof of Theorem 1.11.

3. The free group algebra and its ideals

We gather some known results and some new results about the free group algebra $\mathcal{A} = K[\boldsymbol{F}]$ and its (as always in this text, right) ideals, and more generally the free right $\mathcal{A}$-module $\mathcal{A}^m$ and its submodules. Although we assume throughout this paper that K is a fixed finite field, most results of the current section apply to an arbitrary field (not necessarily finite).

The starting point of the story is a paper of Cohn [1964] and a paper of Lewin [1969] (see note 1) which prove that $\mathcal{A}$ is a *free ideal ring*, in the following sense:

Theorem 3.1 [Cohn 1964; Lewin 1969]. *Every ideal $I \leq \mathcal{A}$ is a free $\mathcal{A}$-module. More generally, every submodule of a free $\mathcal{A}$-module is free. Every free $\mathcal{A}$-module M has a unique rank: all bases of M have the same cardinality.*

See [Hog-Angeloni 1990; Rosenmann and Rosset 1994; Rosenmann 1993] for additional proofs of this result.

There are two main new results in Section 3. In Theorem 3.8 below it is shown that if an ideal $I \leq_T \mathcal{A}$ has a generating set supported on some finite subtree T of $\mathrm{Cay}(\boldsymbol{F}, B)$, then it also admits a basis supported on T. Our analysis also leads to Corollary 3.9: the order of contribution of every ideal $I \leq_{[1,w]} \mathcal{A}$ with $w - 1 \in I$ to the summation (2-5) of $\mathbb{E}_w[\mathrm{fix}]$ is given by its rank.

Recall that $E = \{e_1, \ldots, e_m\}$ is a basis of the free right module $\mathcal{A}^m$, that the elements of $\mathcal{A}^m$ are K-linear combinations of *monomials* $\{ez\}_{e \in E, z \in \boldsymbol{F}}$, and that we identify these monomials with the vertices of m disjoint copies $\mathcal{C}_1, \ldots, \mathcal{C}_m$ of $\mathrm{Cay}(\boldsymbol{F}, B)$. Let $\mathbb{T} = T_1 \cup \cdots \cup T_m$ be a union of m finite, possibly empty, subtrees $T_i \subset \mathcal{C}_i$, and let $M \leq_{\mathbb{T}} \mathcal{A}^m$ be a submodule generated on $\mathbb{T}$. In order to study M, we expose the vertices of $\mathbb{T}$ one-by-one and with them the elements of M which are supported on the already-exposed vertices. Denote by v_t the vertex exposed in step t, where $t = 1, \ldots, \#\mathrm{vert}(\mathbb{T})$, and let M_t denote the submodule generated by $M|_{\{v_1, \ldots, v_t\}}$, so

$$(0) = M_0 \leq M_1 \leq \cdots \leq M_{\#\mathrm{vert}(\mathbb{T})} = M.$$

The order by which we expose the vertices of $\mathbb{T}$ should have the property that as often as possible, we expose neighbors of already-exposed vertices. Formally, it should be the restriction to $\mathbb{T}$ of a full order on the vertices of $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m$ which abides to the following assumption.

Definition 3.2 (exploration). We call a full order $\leq$ on the vertices of $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m$ an *exploration* if it is an enumeration of the vertices (so every vertex has finitely many smaller vertices), and every vertex is either

(1) a neighbor of a smaller vertex, or

(2) the smallest vertex in some $\mathcal{C}_i$.

An order on a collection $\mathbb{T} = T_1 \cup \cdots \cup T_m$ of (possibly empty) subtrees $T_i \subseteq \mathcal{C}_i$ is called an *exploration* if it is the restriction of an exploration of $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m$.

Note that an order on $\mathbb{T}$ is an exploration if and only if it is an enumeration of the vertices of $\mathbb{T}$ which satisfies that every vertex is either a neighbor of a smaller vertex of $\mathbb{T}$ or the first vertex visited in some T_i.

Given a finite $\mathbb{T}$ and $M \leq_{\mathbb{T}} \mathcal{A}^m$ as above, every step is either free, forced or a coincidence, according to the following conventions.[5] Assume first that v_t is a neighbor of an already-exposed vertex u, and that the edge from u to v_t is labeled by $b \in B \cup B^{-1} = \{b_1^{\pm 1}, \dots, b_r^{\pm 1}\}$:

$$u \xrightarrow{b} v_t. \tag{3-1}$$

Denote by D_b^t the set of already-exposed vertices with an outgoing b-edge leading to another already-exposed vertex. This set should include the vertex u. If $M|_{D_b^t}$ contains an element with u in its support, we say the t-th step is *forced*. If v_t is the first vertex we expose in some T_i, the t-th step is not forced. If a step is not forced, it is a *coincidence* if there is an element of $M|_{\{v_1,\dots,v_t\}}$ with v_t in its support, and otherwise it is *free*.

Lemma 3.3. *Let $\mathbb{T}$ and $M \leq_{\mathbb{T}} \mathcal{A}^m$ be as above and let $v_1, v_2, \dots$ be an exploration of* vert($\mathbb{T}$). *Then step t in the exposure of M along $\mathbb{T}$ is*

$$\begin{aligned} \text{forced} &\iff M_{t-1} = M_t \text{ and } M|_{\{v_1,\dots,v_{t-1}\}} \lneqq M|_{\{v_1,\dots,v_t\}}, \\ \text{free} &\iff M_{t-1} = M_t \text{ and } M|_{\{v_1,\dots,v_{t-1}\}} = M|_{\{v_1,\dots,v_t\}}, \\ \text{a coincidence} &\iff M_{t-1} \lneqq M_t. \end{aligned}$$

If step t is a coincidence and f is an element of $M|_{\{v_1,\dots,v_t\}}$ with v_t in its support, then M_t is generated by M_{t-1} and f.

Of course, if $M_{t-1} \lneqq M_t$, then, in particular, $M|_{\{v_1,\dots,v_{t-1}\}} \lneqq M|_{\{v_1,\dots,v_t\}}$.

Proof. First assume step t is forced. There is some $f \in M|_{D_b^t}$ with u in its support, and then $f.b \in M|_{\{v_1,\dots,v_t\}} \setminus M|_{\{v_1,\dots,v_{t-1}\}}$. Yet $f.b \in M_{t-1}$ and any other element of $M|_{\{v_1,\dots,v_t\}}$, by subtracting a suitable K-multiple of $f.b$, becomes an element of M_{t-1}. Hence $M_{t-1} = M_t$.

If the step is free, then $M|_{\{v_1,\dots,v_{t-1}\}} = M|_{\{v_1,\dots,v_t\}}$ by definition, and so $M_{t-1} = M_t$.

Finally, assume that step t is a coincidence. Fix $N \geq t$, and consider (row) vectors $u_1, \dots, u_{t-1} \in V_N = K^N$ with dependencies corresponding *exactly* to the elements of $M|_{\{v_1,\dots,v_{t-1}\}}$, namely, $\sum_{i=1}^{t-1} \alpha_i u_i = 0$ if and only if $\sum_{i=1}^{t-1} \alpha_i v_i \in M$. Let $u_t \in V_N$ be some vector which is *linearly independent* of $u_1, \dots, u_{t-1}$. For every $b \in B$, there is an element $g_b \in \mathrm{GL}(V_N)$ with $u.g_b = u'$ for every b-edge (u, u') with $u, u' \in \{u_1, \dots, u_t\}$ (here we rely on that the step is not forced). As in the proof of Lemma 2.1, these g_b's determine a K-algebra homomorphism $\varphi : \mathcal{A} \to \mathrm{End}(V_N)$. This φ gives V_N a structure of an $\mathcal{A}$-module. For every $e \in E$ with T_e already visited, pick an arbitrary $v_e \in T_e \cap \{v_1, \dots, v_t\}$. Then these monomials $\{v_e\}$ form a subbasis of the free $\mathcal{A}$-module $\mathcal{A}^m$, and there is a homomorphism of $\mathcal{A}$-modules $\psi : \mathcal{A}^m \to V$ mapping v_e to u_e. By design, the linear dependencies among $u_1, \dots, u_t$ correspond precisely to the

[5] This terminology is inspired by [Eberhard and Jezernik 2022], which, in turn, was inspired by earlier works dealing with random Schreier graphs of symmetric groups (see, for example, [Broder and Shamir 1987]). The analog in [Eberhard and Jezernik 2022] of our free step is a free step which is not a coincidence, and the analog in the same article of our coincidence is a free step which is also a coincidence.

elements of $\ker\psi$ supported on $\{v_1,\dots,v_t\}$. As u_t is independent of the rest, we get that

$$M_{t-1}\le\ker\psi\quad\text{yet}\quad M_t\not\le\ker\psi,$$

proving that $M_{t-1}\lneqq M_t$.

If step t is a coincidence and $f\in M|_{\{v_1,\dots,v_t\}}$ has v_t in its support, then any other element $g\in M_{\{v_1,\dots,v_t\}}$ satisfies that $g-\alpha f\in M|_{\{v_1,\dots,v_{t-1}\}}$ for some $\alpha=\alpha(g)\in K$. Hence the final part of the statement of the lemma follows. □

Lemma 3.4. *Let $\mathbb{T}$ and $M\le_{\mathbb{T}}\mathcal{A}^m$ be as above. In every exposure process of M along $\mathbb{T}$ as above, the number of coincidences is the same: it does not depend on the order of exposure (as long as it is a valid exploration à la Definition 3.2).*

Proof. Similarly to the definition of $d^{\mathbb{T}}(M)$ and $d_b^{\mathbb{T}}(M)$ from Section 2, let $d^t\stackrel{\text{def}}{=}\dim_K(M|_{\{v_1,\dots,v_t\}})$ and $d_b^t\stackrel{\text{def}}{=}\dim_K(M|_{D_b^t})$. Obviously, $d^0=d_b^0=0$. We now trace how d^t and $\sum_{b\in B}d_b^t$ change with t, depending on the three types of steps defined above. According to the definitions and to Lemma 3.3:

- In a forced step, both d^t and $\sum_b d_b^t$ increase by one (compared to d^{t-1} and $\sum_b d_b^{t-1}$, respectively).
- In a free step, both d^t and $\sum_b d_b^t$ do not change.
- In a coincidence, d^t increases by one, while $\sum_b d_b^t$ does not change.

Therefore, the difference $d^{\mathbb{T}}(M)-\sum_b d_b^{\mathbb{T}}(M)$, which is, of course, independent of the order of exposure, is equal to the number of coincidences. □

The proof of Lemma 3.4 actually shows that the number of forced and free steps is also independent of the order of exposure, but that is not as useful. The proof also gives the following.

Corollary 3.5. *Consider the expression (2-7) giving $\mathbb{E}_w[\tilde{\mathcal{B}}]$ as a sum over submodules $M\le_{\mathbb{W}}\mathcal{A}^m$ with $M\supseteq\mathrm{EQ}_{\mathcal{B},w}$. The summand corresponding to such a submodule M is*

$$(q^N)^{m-\#\text{coincidences}}\left(1+O\left(\frac{1}{q^N}\right)\right),$$

where we count coincidences in an exposure process of M along $\mathbb{W}$.

Proof. The numerator in the summand corresponding to M in (2-7) is

$$\mathfrak{indep}_{m(|w|+1)-d^{\mathbb{W}}(M)}(V_N)=(q^N)^{m(|w|+1)-d^{\mathbb{W}}(M)}\left(1+O\left(\frac{1}{q^N}\right)\right).$$

The denominator is

$$\prod_b\mathfrak{indep}_{e_b(\mathbb{W})-d_b^{\mathbb{W}}(M)}(V_N)=(q^N)^{\sum_b(e_b(\mathbb{W})-d_b^{\mathbb{W}}(M))}\left(1+O\left(\frac{1}{q^N}\right)\right).$$

The result follows as $\sum_b e_b(\mathbb{W})=m|w|$ and as $d^{\mathbb{W}}(M)-\sum_b d_b^{\mathbb{W}}(M)$ is equal to the number of coincidences. □

Next, we show that the number of coincidences is identical to the rank of the module M. The proof relies on the main theorem of [Lewin 1969], which makes use of the following notion.

Definition 3.6. A *Schreier transversal* of a submodule $M \leq \mathcal{A}^m$ is a set ST of monomials of $\mathcal{A}^m$ which satisfies

(i) ST is closed under prefixes: if $ez \in \mathrm{ST}$ with $e \in E$ and $1 \neq z \in \boldsymbol{F}$, and $b \in B \cup B^{-1}$ is the last letter of z, then $ezb^{-1} \in \mathrm{ST}$, and

(ii) the linear span $\mathrm{span}_K(\mathrm{ST})$ of ST contains exactly one representative of every coset of $\mathcal{A}^m/M$.

It is not hard to show that every $M \leq \mathcal{A}^m$ admits Schreier transversals — see [Lewin 1969, pp. 456–457] for an argument as well as for a concrete construction. A Schreier transversal ST consists of the vertices in a collection of (possibly infinite) subtrees, one in $\mathcal{C}_i$ for every $i = 1, \ldots, m$. The main theorem of [Lewin 1969] is that one may construct a basis for M which is, roughly, in one-to-one correspondence with the outgoing directed edges from ST to its complement. Although a version of this theorem holds for any submodule of any free $\mathcal{A}$-module, we only need the case of finitely generated $\mathcal{A}$-modules.

Theorem 3.7 [Lewin 1969, Theorem 1]. *Let $M \leq \mathcal{A}^m$ be a submodule, and let* ST *be a Schreier transversal of M. For every $f \in \mathcal{A}^m$, denote by $\phi(f)$ the representative of $f + M$ in* $\mathrm{span}_K(\mathrm{ST})$. *Then the set*

$$\{ezb - \phi(ezb) \mid ez \in \mathrm{ST},\ b \in B,\ ezb \notin \mathrm{ST}\} \cup \{e - \phi(e) \mid e \in E \setminus \mathrm{ST}\} \tag{3-2}$$

is a basis for M (as a free $\mathcal{A}$-module).

We stress that in (3-2), b is a proper basis element and not the inverse of one.

Theorem 3.8. *Let $\mathbb{T} = T_1 \cup \cdots \cup T_m$ be a collection of finite, possibly empty, subtrees $T_i \subset \mathcal{C}_i$ and assume that $M \leq_{\mathbb{T}} \mathcal{A}^m$. Then the number of coincidences in an exposure of M along $\mathbb{T}$ is* ***equal*** *to* $\mathrm{rk}\, M$.

Moreover, M admits a basis supported on $\mathbb{T}$. In fact, every set of elements $f_1, \ldots, f_{\mathrm{rk}\, M}$ supported on $\mathbb{T}$ with the leading vertex[6] *of f_i being the monomial exposed in the i-th coincidence is a basis of M.*

Proof. Let $s = \mathrm{rk}\, M$. Let ST be a Schreier transversal for M. Then the basis (3-2) contains s elements. Let $\mathbb{S}$ be the smallest collection of finite subtrees (one in each $\mathcal{C}_i$) which contain the whole support of these s basis elements. Note that $\mathbb{S}$ contains exactly s vertices (monomials) outside ST, and all these vertices are either leaves or isolated in $\mathbb{S}$ (namely, these are vertices of degree 1 or 0 in $\mathbb{S}$). Consider an exposure process of M along $\mathbb{S}$ according to some exploration such that the vertices of $\mathbb{S} \cap \mathrm{ST}$ are exposed first and only then the remaining s vertices. Because there is no nonzero element of M supported on ST, the first $|\mathbb{S}| - s$ steps are all free.

We claim that the remaining s steps are all coincidences. Indeed,

$$(0) = M_{|\mathbb{S}|-s} \leq M_{|\mathbb{S}|-s+1} \leq \cdots \leq M_{|\mathbb{S}|-1} \leq M_{|\mathbb{S}|} = M.$$

For $i = 1, \ldots, s$, let $f_i \in M_{|\mathbb{S}|-s+i}$ be the basis element from (3-2) with the vertex exposed in step $|\mathbb{S}|-s+i$ in its support. Clearly, $f_i \in M_{|\mathbb{S}|-s+i}$. By induction, $M_{|\mathbb{S}|-s+i} = (f_1, \ldots, f_i)$. Indeed, $M_{|\mathbb{S}|-s+1} = (f_1)$, and

[6] The leading vertex of $f \in M$ is the last vertex in the support of f to be exposed in the exploration process on $\mathbb{T}$.

if $M_{|\mathbb{S}|-s+i-1} = (f_1, \dots, f_{i-1})$ then either step i is a coincidence and then $M_{|\mathbb{S}|-s+i} = (M_{|\mathbb{S}|-s+i-1}, f_i)$ by Lemma 3.3, or step i is not a coincidence and then $M_{|\mathbb{S}|-s+i} = M_{|\mathbb{S}|-s+i-1}$. But $\{f_1, \dots, f_s\}$ is a basis by Theorem 3.7, so $f_i \notin (f_1, \dots, f_{i-1}) = M_{|\mathbb{S}|-s+i-1}$. We conclude that $M_{|\mathbb{S}|-s+i-1} \lneqq M_{|\mathbb{S}|-s+i}$ so all these s steps are indeed coincidences by Lemma 3.3.

Now consider the collection of finite trees $\mathbb{U}$, which is the collection of smallest subtrees (one in each $\mathcal{C}_i$) which contains both $\mathbb{S}$ and the given $\mathbb{T}$. Expose M along $\mathbb{U}$ by two different explorations. In the first order, expose $\mathbb{S}$ first and then the remaining vertices of $\mathbb{U}$. There are exactly s coincidences: after we exposed all of $\mathbb{S}$, we have $M_{|\mathbb{S}|} = M$, so no more coincidences are possible, by Lemma 3.3. By Lemma 3.4, s is also the number of coincidences when we first expose $\mathbb{T}$ and then the remaining vertices of $\mathbb{U} \setminus \mathbb{T}$. But again, because M is generated on $\mathbb{T}$, we have $M|_{\mathbb{T}} = M$ and there are no more coincidences after exposing $\mathbb{T}$. This shows there are exactly $s = \operatorname{rk} M$ coincidences in an exposure of M along $\mathbb{T}$.

For the second statement, assume that $M \le_{\mathbb{T}} \mathcal{A}^m$ and consider an exposure of M along $\mathbb{T}$. If step t is a coincidence, then by Lemma 3.3, $M_t = (M_{t-1}, f_t)$ where $f_t \in M|_{\{v_1,\dots,v_t\}}$ with v_t in its support. Hence $M = (f_{t_1}, \dots, f_{t_s})$ where $t_1, \dots, t_s$ are the s coincidences. But every set of size $s = \operatorname{rk} M$ which generates M is a basis [Cohn 1964, Proposition 2.2]. □

From Theorem 3.8 and Corollary 3.5 we immediately obtain that the order of contribution of a given ideal to $\mathbb{E}_w[\mathrm{fix}]$ is given by its rank:

Corollary 3.9. *Consider the expression (2-7) giving $\mathbb{E}_w[\tilde{\mathcal{B}}]$ as a sum over submodules $M \le_{\mathbb{W}} \mathcal{A}^m$ with $M \supseteq \mathrm{EQ}_{\mathcal{B},w}$. The summand corresponding to such a submodule M is*

$$(q^N)^{m-\operatorname{rk} M}\left(1 + O\left(\frac{1}{q^N}\right)\right).$$

In Section 4B we show that there are no submodules of rank $< m$ containing $\mathrm{EQ}_{\mathcal{B},w}$, and so $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$, the limit from Theorem 1.12, is equal to the number of rank-m submodules supported on $\mathbb{W}$ and containing $\mathrm{EQ}_{\mathcal{B},w}$. Using Corollary 3.10, one can show that the restriction to submodules supported on $\mathbb{W}$ is redundant — we elaborate in Section 4.

Recall that Definition 1.5 introduced $\pi_q(w)$ and $\mathrm{Crit}_q(w)$ for every $w \in \boldsymbol{F}$. Theorem 3.8 can also be used to show that the set $\mathrm{Crit}_q(w)$ is always finite. If N is a free $\mathcal{A}$-module and $L \le N$ a submodule (and therefore free as well), we say that L is a free factor of N if some basis (and hence every basis) of L can be extended to a basis of N.

Corollary 3.10. *Let $M \le N \le \mathcal{A}^m$ be two finitely generated submodules of $\mathcal{A}^m$, and assume that there is no intermediate submodule which is a proper free factor of N.*[7] *Namely, if $M \le L \le N$ and L is a free factor of N then $L = N$. If $M \le_{\mathbb{T}} \mathcal{A}^m$ with $\mathbb{T}$ a union of subtrees as above, then $N \le_{\mathbb{T}} \mathcal{A}^m$.*

Proof. Take a collection $\mathbb{S}$ of subtrees which contains $\mathbb{T}$ and such that $N \le_{\mathbb{S}} \mathcal{A}^m$. Expose N along $\mathbb{S}$ according to some exploration which first exposes $\mathbb{T}$ and then the remaining vertices. Let $N_{|\mathbb{T}|} = (N|_{\mathbb{T}})$

[7] In analogy with subgroups of the free group $\boldsymbol{F}$, one may say that N is an *algebraic* extension of M — see, e.g., [Puder and Parzanchevski 2015, Definition 2.1].

denote the submodule of N generated by the elements of N supported on $\mathbb{T}$. Clearly, $M \le N_{|\mathbb{T}|}$, and using Theorem 3.8 to construct a basis for N from the coincidences of this exposure process, we get that $N_{|\mathbb{T}|}$ is a free factor of N. By assumption we therefore have $N_{|\mathbb{T}|} = N$, so $N \le_{\mathbb{T}} \mathcal{A}^m$. □

In the following corollary we use the fact that K is finite. For example, the element $xyx^{-1}y^{-1} - 1 \in \mathcal{A}$ has critical ideals $\{(\alpha x - 1, \beta y - 1) \mid \alpha, \beta \in K^*\}$, which is an infinite set if K is infinite. For a general element $f \in \mathcal{A}$, we say that an ideal $I \le \mathcal{A}$ is *critical* for f if it contains f as an imprimitive element, and it has minimal rank among all such ideals.

Corollary 3.11. *Let $f \in \mathcal{A}$, and suppose that the subtree $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ supports f. Then any critical ideal of f is generated on T. In particular, $\mathrm{Crit}_q(w)$ is finite for every word $w \in \boldsymbol{F}$ and every prime power q.*

Proof. Assume that $I \le \mathcal{A}$ is critical for f, namely, that it is an ideal of minimal rank which contains f as an imprimitive element. Assume that $f \in J \le I$ and that J is a free factor of I. In particular, $\mathrm{rk}\, J \le \mathrm{rk}\, I$. If f is primitive in J, it is also primitive in I, which is impossible. So f is imprimitive in J. But I is critical for f, and so $\mathrm{rk}\, J = \mathrm{rk}\, I$ and $J = I$. Therefore the assumption of Corollary 3.10 applies to $(f) \le I$, and for every finite subtree $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ supporting f, we have $I \le_T \mathcal{A}$. For every $f \in \mathcal{A}$ we may take T finite, and if K is finite, there are only finitely many ideals supported on T. □

3A. ***Properties of the q-primitivity rank.*** We prove some basic properties of the q-primitivity rank of words. Let H be a subgroup of the free group $\boldsymbol{F}$. We associate to H two (right) ideals of interest. The first is its augmentation ideal $I_H \le K[H]$, defined as the kernel of the augmentation map $\varepsilon_H : K[H] \to K$ where $\varepsilon_H\big(\sum_{h\in H} \alpha_h h\big) = \sum_{h\in H} \alpha_h$. If $\{h_\beta\}_{\beta\in B}$ is a basis for H then $\{h_\beta - 1\}_{\beta\in B}$ is a basis for I_H [Cohen 1972, Proposition 4.8], and in particular $\mathrm{rk}\, I_H = \mathrm{rk}\, H$. The second, when considering H as a subgroup of $\boldsymbol{F}$, is the (right) ideal J_H of $\mathcal{A} = K[\boldsymbol{F}]$ generated by $\{h - 1\}_{h\in H}$. The following proposition also follows from [Cohen 1972, Chapter 4], but as it is not stated there explicitly, we add a short proof.

Proposition 3.12. *If $\{h_\beta\}_{\beta\in B}$ is a basis for H then $\{h_\beta - 1\}_{\beta\in B}$ is a basis for J_H. In particular, $\mathrm{rk}\, J_H = \mathrm{rk}\, H$.*

Proof. Since $\{h_\beta - 1\}_{\beta\in B}$ already generates I_H in $K[H]$, it generates $h - 1$ for any $h \in H$, and is thus a generating set for J_H. Let T be a right transversal for H in $\boldsymbol{F}$ (i.e., a set of representatives of the right cosets of H). Then for every $t \in T$ the set $K[H]t$ of elements of $\mathcal{A}$ supported on the coset Ht forms a left $K[H]$-module, and the group algebra $\mathcal{A}$ admits a left $K[H]$-module decomposition $\mathcal{A} = \bigoplus_{t\in T} K[H]t$. Let $P_{Ht} : \mathcal{A} \to K[H]t$ be the projections induced by this decomposition. Suppose now that there is a relation $\sum_{\beta\in B}(h_\beta - 1)a_\beta = 0$ for some coefficients $\{a_\beta\}_{\beta\in B}$ in $\mathcal{A}$. For every $t \in T$, applying the left $K[H]$-module map P_{Ht} to both sides yields the relation $\sum_{\beta\in B}(h_\beta - 1)P_{Ht}(a_\beta) = 0$, and multiplying by t^{-1} gives $\sum_{\beta\in B}(h_\beta - 1)(P_{Ht}(a_\beta)t^{-1}) = 0$. Since $P_{Ht}(a_\beta)t^{-1} \in K[H]$ and $\{h_\beta - 1\}_{\beta\in B}$ is a basis (for I_H), we deduce that $P_{Ht}(a_\beta) = 0$ for every $\beta \in B$. Thus, $a_\beta = \sum_{t\in T} P_{Ht}(a_\beta) = 0$ for every $\beta \in B$. □

Proposition 3.13. *Let $H \le \boldsymbol{F}$ be finitely generated and let $w \in \boldsymbol{F}$. If $w - 1$ is primitive in J_H then w is primitive in H.*

Proof. Assume that $w-1$ is primitive in J_H. As $w-1 \in J_H$, by [Cohen 1972, Lemma 4.1], w lies in H. Fix a basis $h_1, h_2, \dots, h_k$ for H. Then $\{h_i - 1\}_{i=1}^k$ is a basis for I_H and $w-1 \in I_H$, so we can write (uniquely) $w - 1 = \sum_{i=1}^k (h_i - 1)a_i$ for some coefficients $a_i \in K[H]$. By a theorem of Umirbaev[8] [1994, Corollary on page 184], to deduce that w is primitive in H it is enough to show that the coefficients $\{a_i\}_{i=1}^k$ form a left-invertible column in the sense that there exist $u_1, u_2, \dots, u_k \in K[H]$ such that $\sum_{i=1}^k u_i a_i = 1$. Since $w-1$ is primitive in J_H, there exist some elements $f_2, \dots, f_k \in J_H$ completing $w-1$ to a basis of J_H. By Proposition 3.12, $\{h_i - 1\}_{i=1}^k$ is, too, a basis for J_H. Let $C \in M_{kk}(\mathcal{A})$ be a change-of-basis matrix satisfying $(h_1 - 1, h_2 - 1, \dots, h_k - 1)C = (w-1, f_2, \dots, f_k)$, where by uniqueness of presenting $w-1$ in the basis $\{h_i - 1\}_{i=1}^k$ the first column of C is

$$\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}.$$

As one can also change basis in the other direction, there exists some $D \in M_{kk}(\mathcal{A})$ such that

$$(h_1 - 1, h_2 - 1, \dots, h_k - 1) = (w-1, f_2, \dots, f_k)D.$$

Thus, $(w-1, f_2, \dots, f_k)DC = (w-1, f_2, \dots, f_k)$, which by the uniqueness of presentation implies that DC is the identity matrix. In particular, the first row of D which we denote by $(d_1, d_2, \dots, d_k)$ is a left inverse to the first column of C in the sense that

$$\sum_{i=1}^k d_i a_i = 1. \tag{3-3}$$

We next show that the elements $\{d_i\}_{i=1}^k$ can be replaced by elements $\{u_i\}_{i=1}^k$ lying in $K[H]$. Let T be a left transversal for H in $\boldsymbol{F}$. Then as a right $K[H]$-module, $\mathcal{A}$ decomposes as $\mathcal{A} = \bigoplus_{t \in T} tK[H]$. Denote by P_{tH} the projection onto the summand corresponding to t. Then applying the right $K[H]$-module map P_H to (3-3) gives $\sum_{i=1}^k P_H(d_i)a_i = 1$. We finish by letting $u_i = P_H(d_i)$. □

Lemma 3.14. *Let $J \leq \mathcal{A}$ be an ideal and $f \in J$ a primitive element. Then f is primitive in every intermediate ideal $f \in I \leq J$.*

Proof. Since f is primitive in J we can write $J = (f) \oplus M$ for some ideal M. We claim that $I = (f) \oplus (M \cap I)$. The directness is obvious (M already intersects (f) trivially). It remains to show that I is indeed the sum of the two summands. Let $a \in I$. Then $a \in J$ and thus can be decomposed as $a = a_1 + m$ where $a_1 \in (f)$ and $m \in M$. But then $m = a - a_1 \in I$ and so $m \in M \cap I$. □

In the following corollary we do not assume that H is finitely generated.

Corollary 3.15. *Let $H \leq \boldsymbol{F}$ and let $w \in \boldsymbol{F}$. Then w is primitive in H if and only if $w-1$ is primitive in J_H.*

[8] Umirbaev's result is actually stated for free group rings over the integers. However, the proof uses no specific properties of $\mathbb{Z}$ and hence also applies, mutatis mutandis, to the field K.

Proof. One implication is immediate from Proposition 3.12. For the other implication, suppose that $w-1$ is primitive in J_H. Let S be a basis for H. When writing $w-1$ in the basis $\{s-1\}_{s\in S}$ of J_H, only finitely many basis elements appear, so there exist some $h_1,\dots,h_k \in S$ and $a_1,\dots,a_k \in \mathcal{A}$ such that $w-1=\sum_{i=1}^{k}(h_i-1)a_i$. Let $H'=\langle h_1,h_2,\dots,h_k\rangle$. Then $w-1$ lies in $J_{H'}$ and by Lemma 3.14 it is primitive in it. Since H' is finitely generated, Proposition 3.13 guarantees that w is primitive in H'. Since the relation of being a free factor is transitive and $\langle w\rangle \overset{*}{\leq} H' \overset{*}{\leq} H$ we are done. □

We can now prove Proposition 1.8 stating that for every prime power q, the q-primitivity rank is bounded from above by the ordinary primitivity rank, namely, $\pi_q(w)\leq\pi(w)$ for every $w\in\boldsymbol{F}$.

Proof of Proposition 1.8. Let $w\in\boldsymbol{F}$. The ordinary primitivity rank of a word is a nonnegative integer or ∞. We first deal with two trivial cases: if $\pi(w)=\infty$ then there is nothing to prove, and if $\pi(w)=0$ then $w=1$ and so $w-1=0$ is contained in the rank-0 trivial ideal of $\mathcal{A}$ as an imprimitive element, so $\pi_q(w)=0$ as well. Suppose now that $\pi(w)=k\notin\{0,\infty\}$ and let H be a critical subgroup for w in $\boldsymbol{F}$, i.e., a subgroup of $\boldsymbol{F}$ of rank k containing w as an imprimitive element. The ideal $J_H\leq\mathcal{A}$ contains $w-1$ by its definition as $w\in H$, it contains $w-1$ as an imprimitive element by Corollary 3.15, it has rank $\operatorname{rk} J_H=k$ by Proposition 3.12, and it is a proper ideal of $\mathcal{A}$ since it is contained in the augmentation ideal $I_{\boldsymbol{F}}\subsetneq\mathcal{A}$. We conclude that $\pi_q(w)\leq k=\pi(w)$. □

If $w\in\boldsymbol{F}$ is primitive, then (analogously to Lemma 3.14), w is primitive in any subgroup of $\boldsymbol{F}$ containing it (see, e.g., [Puder 2014, Claim 2.5]). In particular, it has primitivity rank $\pi(w)=\infty$. Furthermore, any imprimitive word $w\in\boldsymbol{F}$ must have $\pi(w)\leq\operatorname{rk}\boldsymbol{F}$ since it is already not primitive in $\boldsymbol{F}$. Thus, the primitivity rank of words in $\boldsymbol{F}$ takes values in $\{0,1,2,\dots,\operatorname{rk}\boldsymbol{F}\}\cup\{\infty\}$. We next show that analogous statements hold when π is replaced with π_q.

Proposition 3.16. *For every $w\in\boldsymbol{F}$ and prime power q, $\pi_q(w)=\infty$ if and only if w is primitive in $\boldsymbol{F}$.*

Proof. If $\pi_q(w)=\infty$ then $w-1$ must be primitive in $J_{\boldsymbol{F}}$, which implies by Proposition 3.13 that w is primitive in $\boldsymbol{F}$. Conversely, let $w\in\boldsymbol{F}$ be primitive. Then there exists some automorphism $\psi\in\operatorname{Aut}\boldsymbol{F}$ such that $\psi(w)=b_1$ (recall that $\{b_1,\dots,b_r\}$ is our fixed basis of $\boldsymbol{F}$). The automorphism ψ naturally extends (linearly) to an automorphism of the group ring $\psi:\mathcal{A}\to\mathcal{A}$. Since ψ maps ideals to ideals and bases to bases, it is enough to show that $\psi(w-1)=b_1-1$ is primitive in every ideal containing it. Suppose it is not, and let I be a critical ideal for b_1-1. By Corollary 3.11, I is generated on $T=\{1,b_1\}$. Let $f\in I|_T$. Then $f=\beta b_1-\alpha$ for some $\alpha,\beta\in K$. By the definition of a critical ideal, I is a proper ideal and so α,β must be equal because their difference lies in I:

$$\beta-\alpha=f-(b_1-1)\beta\in I.$$

Thus, $I|_T=\operatorname{span}_K\{b_1-1\}$ and since I is supported on T, I must be the right principal ideal $I=(b_1-1)$ in which b_1-1 is primitive, a contradiction. □

Corollary 3.17. *For every $w\in\boldsymbol{F}$ and every prime power q, $\pi_q(w)\in\{0,1,2,\dots,\operatorname{rk}\boldsymbol{F}\}\cup\{\infty\}$.*

Proof. Let $w \in \boldsymbol{F}$. If w is primitive in $\boldsymbol{F}$ then by Proposition 3.16 $\pi_q(w) = \infty$. Otherwise, by Corollary 3.15, $w-1$ is already imprimitive in $J_{\boldsymbol{F}}$ and so $\pi_q(w) \le \mathrm{rk}\, J_{\boldsymbol{F}} = \mathrm{rk}\, \boldsymbol{F}$. □

4. Powers and the limit of expected values of stable functions

We prove Theorems 1.2 and 1.12: if $w \neq 1$, then $\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}]$ and, more generally, $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$, exist. If we write $w = u^d$ with u a nonpower and $d \ge 1$, then the limit depends only on d, and, in particular, $\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}]$ is equal to the number of monic divisors of the polynomial $x^d - 1 \in K[x]$.

As mentioned in Section 1, a special case of Theorem 1.12, which includes Theorem 1.2, first appeared in [Ernst-West 2019] and, independently, in [Eberhard and Jezernik 2022]. Here, we prove the full version of the theorem while following the strategy from [Eberhard and Jezernik 2022], which is more elegant than the one in [Ernst-West 2019]. We use here slightly different notions and give more details than in [Eberhard and Jezernik 2022]. As the proof is subtle, and for the sake of readability, we first describe the proof of the special case which is Theorem 1.2.

4A. $\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{w}}[\mathbf{fix}]$ *and the proof of Theorem 1.2.* From (2-5) and Corollary 3.9 it follows that

$$\mathbb{E}_w[\mathrm{fix}] = \sum_{I \le_{[1,w]} \mathcal{A}: I \ni w-1} (q^N)^{1-\mathrm{rk}\, I} \left(1 + O\left(\frac{1}{q^N}\right)\right).$$

Clearly, as $w \neq 1$, an ideal containing $w-1$ has rank at least 1. So

$$\mathbb{E}_w[\mathrm{fix}] = |\{I \le_{[1,w]} \mathcal{A} \mid I \ni w-1 \text{ and } \mathrm{rk}\, I = 1\}| + O\left(\frac{1}{q^N}\right). \tag{4-1}$$

By Definition 1.5, the only noncritical rank-1 ideals containing $w-1$ are $(w-1)$ and (1), which are both generated on $[1, w]$. Any other rank-1 ideal containing $w-1$ is critical, and Corollary 3.11 guarantees that such ideals are supported on $[1, w]$. We obtain that

$$\mathbb{E}_w[\mathrm{fix}] = |\{I \le \mathcal{A} \mid I \ni w-1 \text{ and } \mathrm{rk}\, I = 1\}| + O\left(\frac{1}{q^N}\right). \tag{4-2}$$

This proves:

Corollary 4.1. *Conjecture 1.6 holds in the case $\pi_q(w) = 1$. Namely, in this case*

$$\mathbb{E}_w[\mathrm{fix}] = 2 + |\mathrm{Crit}_q(w)| + O\left(\frac{1}{q^N}\right).$$

In order to prove Theorem 1.2, it remains to show that if $w = u^d$ with u a nonpower and $d \ge 1$, then the ideals I in (4-2) are precisely $\{(p(u)) : p \mid x^d - 1 \in K[x],\ p \text{ monic}\}$. First, as any automorphism of $\boldsymbol{F}$ gives rise to an automorphism of $\mathcal{A}$, we may replace w by any element in its $\mathrm{Aut}\,\boldsymbol{F}$-orbit, and, in particular, assume that w is cyclically reduced.

Throughout Section 4, we use the ShortLex order on monomials in $\mathcal{A}^m$ and their finite subsets.

Definition 4.2. Fix an arbitrary full order on the basis E of $\mathcal{A}^m$, say $e_1 < e_2 < \cdots < e_m$. Fix an arbitrary full order on $B \cup B^{-1}$, say $b_1 < b_1^{-1} < b_2 < \cdots < b_r^{-1}$. The *ShortLex* order on the monomials $\{ez\}_{e\in E, z\in \boldsymbol{F}}$ is defined by first comparing the length of z (shorter words are smaller) and using lexicographic order to compare ez with $e'z'$ when $|z| = |z'|$. This order induces a full order on finite sets of monomials by comparing the leading monomial in each set, breaking ties by looking at the second monomials, and so on (the empty set is the smallest of all finite sets of monomials). Finally, we get a preorder on the elements of $\mathcal{A}^m$ by comparing their supports. An element $f \in \mathcal{A}^m$ is called *monic* if the K-coefficient of the leading monomial is 1.

For example, $\alpha e_2 b_1^{-1} b_2 < \beta e_2 b_1^{-1} b_2 + e_3 b_1$ and $e_1 b_1 b_2 < e_1 b_3 b_1 < e_2 b_1 b_2$ (here $\alpha, \beta \in K^*$). This ShortLex order is the same one used in [Rosenmann 1993]. (The order used in [Lewin 1969] is not quite the same: it uses length and then *reverse* lexicographic order, and it also fixes a full order on K resulting in a full order on $\mathcal{A}^m$, rather than a mere preorder.)

Now, let $I \le \mathcal{A}$ be a rank-1 ideal containing $w - 1$. As noted above, I is generated on $[1, w]$. In the notation of Section 3, consider the exposure of I along the subtree $[1, w]$, starting with the monomial 1 and ending with the monomial w. (This happens to be the restriction of ShortLex to $[1, w]$.) We shift the indices of the vertices by one with respect to Section 3, and define $v_0 = 1, \ldots, v_{|w|} = w$. By Theorem 3.8, there is exactly one coincidence is this exposure.[9] In an exposure along a path, a free step is followed by either another free step or by a coincidence, and in this particular path, the last step is not free. Thus, the first nonfree step must be a coincidence, and the following steps must all be forced. Namely, if v_t is exposed in a coincidence, $t \in \{0, 1, \ldots, |w|\}$, then $v_0, \ldots, v_{t-1}$ are free steps and $v_{t+1}, \ldots, v_{|w|}$ are forced. Denote by $f_I \in I$ the monic element supported on $[1, w]$ with v_t its leading monomial. By Theorem 3.8, $I = (f_I)$. Thus, the map

$$I \mapsto f_I \tag{4-3}$$

is a one-to-one correspondence.

Lemma 4.3. *The ideals I for which f_I is supported on $\langle u\rangle$ are in one-to-one correspondence with monic polynomials in $K[x]$ dividing $x^d - 1$.*

Proof. Consider the subalgebra $K[\langle u\rangle]$ of $\mathcal{A} = K[\boldsymbol{F}]$, the elements of which are linear combinations of the elements in $\langle u\rangle = \{u^i \mid i \in \mathbb{Z}\}$. For every $z \in \boldsymbol{F}$ and $f \in K[\langle u\rangle]$, if $z \notin \langle u\rangle$ then fz is supported on $\langle u\rangle z$, which is disjoint from $\langle u\rangle$. Thus, if $f \in K[\langle u\rangle]$ and $w - 1 = u^d - 1 \in (f) \overset{\text{def}}{=} f\mathcal{A}$, then $w - 1$ is also an element of $fK[\langle u\rangle]$, the ideal generated by f inside $K[\langle u\rangle]$. Now

$$K[\langle u\rangle] \cong K[\mathbb{Z}] \cong K[x, x^{-1}]$$

is a commutative ring (a principal ideal domain, in fact). If $p \in K[x, x^{-1}]$ satisfies $(p) \ni x^d - 1$, we may assume, by multiplying p by a unit element if need be, that $p \in K[x]$, p monic, and $p \mid x^d - 1$

[9] We remark that to analyze $\lim_{N\to\infty} \mathbb{E}_w[\mathrm{fix}]$, one does not really need to go through Theorem 3.8, nor even consider the rank of ideals. Rather, it is enough to rely on Corollary 3.5.

in $K[x]$. Moreover, such $p \in K[x]$ is determined uniquely by the ideal (p). This completes the proof of the lemma. □

It remains to show that for every I in (4-2), f_I is supported on $\langle u \rangle$.

Lemma 4.4. *The support of f_I contains $v_0 = 1$.*

Proof. Recall that t denotes the coincidence step in the exposure of I along $[1, w]$. If $t = |w|$, then $f_I = w - 1$ and the lemma holds. Assume that $t < |w|$ and that the support of f_I does not contain 1. For every $s \geq t$, the vertex v_s is (exposed in) a nonfree step, and let $f_s \in I$ be the ShortLex-minimal element among all monic elements in I with v_s their leading monomial and which are supported on $\{v_0, \ldots, v_s\}$. (This definition is unambiguous: if $f \neq g$ are two different monic elements with the exact same support, then there is some linear combination $\lambda f + (1-\lambda)g$ which is strictly smaller). In particular, $f_t = f_I$ and $f_{|w|} = w - 1$. Now fix $s \in \{t+1, \ldots, |w|\}$ to be the smallest index for which $v_0 = 1$ is in the support of f_s. There is no element in $I|_{\{v_0,\ldots,v_{s-1}\}}$ with v_0 in its support, because $I|_{\{v_0,\ldots,v_{s-1}\}} = \mathrm{span}_K\{f_t, \ldots, f_{s-1}\}$. Let $b \in B \cup B^{-1}$ denote the label of the edge from v_{s-1} to v_s. As step s is forced, there is some monic $g \in I|_{\{v_1,\ldots,v_{s-1}\}}$ with leading monomial v_{s-1} such that $g.b$ is supported on $\{v_0, \ldots, v_s\}$. This $g.b$ must have v_0 in its support, for if not, $f_s - g.b$ does, and the latter is supported on $\{v_0, \ldots, v_{s-1}\}$. We conclude that the first edge of w must be b^{-1}. As w is cyclically reduced, $s < |w|$.

Now consider the vertex v_{s+1}, and assume the edge from v_s to v_{s+1} is $c \in B \cup B^{-1}$, $c \neq b^{-1}$:

$$v_{s-1} \xrightarrow{b} v_s \xrightarrow{c} v_{s+1}.$$

As $I|_{\{v_0,\ldots,v_s\}} = \mathrm{span}_K\{f_t, \ldots, f_s\}$, every element $g \in I|_{\{v_0,\ldots,v_s\}}$ with leading monomial v_s must have v_0 in its support. But then, g cannot possibly be supported on starting points of c-edges, contradicting the fact that step $s+1$ is also forced. Thus f_I contains v_0 in its support. □

If $t = 0$ then $f_I = 1$ and $I = (1)$. If $t = |w|$, then $f_I = w - 1$ and $I = (w-1)$. So assume from now on that $0 < t < |w|$. Also, denote by $b \in B \cup B^{-1}$ the first letter of w.

Lemma 4.5. *The letter from v_t to v_{t+1} is b, and f_I must be supported on D_b^{t+1}.*

Proof. Assume the edge from v_t to v_{t+1} is c. The step exposing v_{t+1} is forced, but f_I is the only monic element in $I|_{\{v_0,\ldots,v_t\}}$. Thus the corresponding element in D_c^{t+1} must be f_I. As f_I has v_0 in its support, we must have $c = b$. □

Completing the proof of Theorem 1.2. Recall that we now assume that I is a rank-1 ideal containing $w - 1$ with $I \neq (1), (w-1)$, and that we need to show that I is supported on $\langle u \rangle$. As I contains $w - 1$ if and only if it contains $wb - b$, the argument above (and Corollary 3.11) show that $I \leq_{[b,wb]} \mathcal{A}$. Expose I along $[b, wb]$ according to the restriction of ShortLex, and denote the vertices by $v_1, \ldots, v_{|w|+1}$ (so keeping the same labels as before). As f_I is the only monic element in $I|_{\{v_0,\ldots,v_t\}}$, we have that $I|_{\{v_1,\ldots,v_t\}} = 0$, that v_{t+1} is the first (and only) coincidence in the exposure along $[b, wb]$, and that the coincidence is given by $f_I.b$. Clearly, $f_I.b$ has v_1 in its support. If b' is the second letter of w, then the same argument

as in Lemma 4.5 shows that $f_I.b$ is supported on vertices with an outgoing b'-edge in $[b, wb]$, and that the edge from v_{t+1} to v_{t+2} is b'.

By iterating the same argument we get that for every prefix w' of w, $f_I.w'$ is supported on $[1, w^2]$. Moreover, the direction in which one can read a prefix of w from some v_j in the support of f_I along $[1, w^2]$ is necessarily forward: if it goes backward, then after $\lfloor \frac{j}{2} \rfloor$ step this path would collide with the path reading w coming from v_0 (a letter in $[1, w]$ cannot be equal to its own inverse or to the inverse of the following letter). We obtain that if f_I has some $v_j = z \in \boldsymbol{F}$ in its support, then $zw = wz$, and so z belongs to the centralizer of w in $\boldsymbol{F}$, which is $\langle u \rangle$. This completes the proof. □

Corollary 4.6. *Let $1 \neq w \in \boldsymbol{F}$. Then $\pi_q(w) = 1$ if and only if w is a proper power.*

Proof. Write $w = u^d$ with u a nonpower and $d \geq 1$. The discussion above shows that the rank-1 critical ideals of $w - 1$ are in one-to-one correspondence with the monic divisors of $x^d - 1 \in K[x]$, except for 1 and $x^d - 1$. If $d = 1$, there are no such divisors, and so $\pi_q(w) \geq 2$. If $d \geq 2$, there is at least one such divisor: the polynomial $x - 1$, and so $\pi_q(w) = 1$. □

Recall that if $\lambda \in \mathrm{GL}_1(K) \cong K^*$, then $\tilde{\lambda} : \mathrm{GL}_N(K) \to \mathbb{Z}_{\geq 0}$ counts, for every element $g \in \mathrm{GL}_N(K)$, the number of vectors $v \in V = K^N$ satisfying $v.g = \lambda v$. The same argument given above for $\mathrm{fix} = \tilde{1}$ applies to all $\lambda \in K^*$ and gives the following result.

Corollary 4.7. *Let $\lambda \in \mathrm{GL}_1(K) \cong K^*$, let $1 \neq w \in \boldsymbol{F}$ and write $w = u^d$ with u a nonpower and $d \geq 1$. Then,*

$$\lim_{N \to \infty} \mathbb{E}_w[\tilde{\lambda}] = \big|\{p \in K[x] : p \mid x^d - \lambda \text{ and } p \text{ monic}\}\big|.$$

4B. $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$ *and the proof of Theorem 1.12.* Our next goal is proving Theorem 1.12, which states that for any fixed $\mathcal{B} \in \mathrm{GL}_m(K)$, the limit $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$ exists and depends only on d, where $w = u^d \neq 1$ as before. As in the proof of Theorem 1.2, we may assume that w is cyclically reduced. From (2-7) and Corollary 3.9 it follows that

$$\mathbb{E}_w[\tilde{\mathcal{B}}] = \sum_{M \leq_{\mathbb{W}} \mathcal{A}^m : M \supseteq \mathrm{EQ}_{\mathcal{B},w}} (q^N)^{m - \mathrm{rk}\, M} \left(1 + O\left(\frac{1}{q^N}\right)\right), \tag{4-4}$$

where $\mathbb{W}$ is the union of the paths $[e_i, e_i w] \in \mathcal{C}_i$ for $i = 1, \ldots, m$. Throughout this Section 4B we continue using ShortLex from Definition 4.2 and its restriction to collections of subtrees such as $\mathbb{W}$.

Lemma 4.8. *The smallest rank of a submodule $M \leq_{\mathbb{W}} \mathcal{A}^m$ containing $\mathrm{EQ}_{\mathcal{B},w}$ is m. In particular, $\lim_{N\to\infty} \mathbb{E}_w[\tilde{\mathcal{B}}]$ exists and*

$$\lim_{N \to \infty} \mathbb{E}_w[\tilde{\mathcal{B}}] = |\{M \leq_{\mathbb{W}} \mathcal{A}^m \mid M \supseteq \mathrm{EQ}_{\mathcal{B},w} \text{ and } \mathrm{rk}\, M = m\}|. \tag{4-5}$$

Proof. Let $M \leq_{\mathbb{W}} \mathcal{A}^m$ contain $\mathrm{EQ}_{\mathcal{B},w}$. We expose M along $\mathbb{W}$ in the order induced on $\mathbb{W}$ from ShortLex. So we first expose $e_1, \ldots, e_m$, then, if the first letter of w is $b \in B \cup B^{-1}$, we expose $e_1 b, \ldots, e_m b$, and so on. By definition, the last m steps, where $e_1 w, \ldots, e_m w$ are exposed, are not free: $\mathrm{EQ}_{\mathcal{B},w}$ contains an element supported on $\{e_i w, e_1, \ldots, e_m\}$ with $e_i w$ its leading monomial. We claim that for every $e \in E$,

the first nonfree step in $[e, ew]$ is a coincidence. In particular, there are at least m coincidences and so by Theorem 3.8, $\mathrm{rk}\, M \geq m$.

Indeed, assume that the first nonfree vertex in $[e_i, e_i w]$ is $e_i z$ for some prefix z of w. If $z = 1$, then $e_i z$ is a coincidence by definition. Now assume that $z \neq 1$ and that $b \in B \cup B^{-1}$ is the last letter of z. As $e_i z$ is the first nonfree step in $[e_i, e_i w]$, we have that $e_i z b^{-1}$ was free, so there is no element of $M|_{\mathbb{W}}$ with leading monomial $e_i z b^{-1}$. Between the exposure of $e_i z b^{-1}$ and that of $e_i z$, the vertices exposed do not admit outgoing b-edges (in the already-exposed part of $\mathbb{W}$): these vertices are either $e_j z b^{-1}$ for $j > i$, where the only outgoing edge is headed backwards and cannot be b as w is reduced; or $e_j z$ for $j < i$, where the only outgoing edge is b^{-1}. Thus, when exposing $e_i z$ at step t, the largest monomial in D_b^t is $e_i z b^{-1}$, but as $e_i z b^{-1}$ is free, there are no elements of $M|_{D_b^t} \subseteq M|_{\mathbb{W}}$ with leading monomial $e_i z b^{-1}$. So step t cannot be forced and must be a coincidence. □

The proof of Lemma 4.8 actually shows that a free vertex in $[e, ew]$ cannot be followed by a forced vertex in the same path. As the last vertex in $[e, ew]$ is nonfree, we get the following.

Corollary 4.9. *If $M \leq_{\mathbb{W}} \mathcal{A}^m$ has rank m and contains $\mathrm{EQ}_{\mathcal{B},w}$, then for every $e \in E$, the first nonfree step in $[e, ew]$ is a coincidence, and all later steps in $[e, ew]$ are forced.*

Remark 4.10. It is possible to extend Corollary 3.11 from elements and ideals in $\mathcal{A}$ to subsets and submodules in $\mathcal{A}^m$, and conclude that every rank-m submodule of $\mathcal{A}^m$ containing $\mathrm{EQ}_{\mathcal{B},w}$ is supported on $\mathbb{W}$.

Lemma 4.11. *Assume that $1 \neq w = u^d$ with $d \geq 1$ and u a nonpower. To prove Theorem 1.12, it is enough to show that every submodule $M \leq_{\mathbb{W}} \mathcal{A}^m$ of rank m with $M \supseteq \mathrm{EQ}_{\mathcal{B},w}$ is generated on $\{eu^j\}_{e\in E, j\in\{0,\dots,d\}}$.*

Proof. Assume that every submodule M from (4-5) is generated on $\{eu^j\}_{e\in E, j\in\{0,\dots,d\}}$. Then, as in the proof of Lemma 4.3, these submodules are in one-to-one correspondence with rank-m submodules of $K[\langle u\rangle]^m$ containing $\mathrm{EQ}_{\mathcal{B},w}$ (and generated on $\{eu^j\}_{e\in E, j\in\{0,\dots,d\}}$), where $K[\langle u\rangle]^m$ is the rank-m free module over $K[\langle u\rangle]$. As before, $K[\langle u\rangle] \cong K[\mathbb{Z}] \cong K[x, x^{-1}]$, and the image of $\mathrm{EQ}_{\mathcal{B},w} \subseteq K[\langle u\rangle]^m$ in $K[\mathbb{Z}]^m$ through the corresponding isomorphism does not depend on u but only on d. Hence, the number of submodules in (4-5) does not depend on u, proving Theorem 1.12. □

Remark 4.12. It is quite straightforward to show that every submodule of $K[\langle u\rangle]^m$ containing $\mathrm{EQ}_{\mathcal{B},w}$ must be of rank exactly m: after the first coincidence in each of the m paths, all remaining steps are clearly forced.

Now fix $M \leq_{\mathbb{W}} \mathcal{A}^m$ of rank m containing $\mathrm{EQ}_{\mathcal{B},w}$. For every $f \in M|_{\mathbb{W}}$, denote by $\theta(f)$ the projection of f to the monomials $e_1, \dots, e_m$, so $\theta(f)$ is a K-linear combination of $e_1, \dots, e_m$. For $t = 0, \dots, |\mathbb{W}|$, let

$$\Theta_t \stackrel{\mathrm{def}}{=} \mathrm{span}_K\{\theta(f) \mid f \in M|_{D^t(\mathbb{W})}\} \leq \mathrm{span}_K\{e_1, \dots, e_m\}$$

(recall that $D^t(\mathbb{W})$ is the set of first t monomials exposed in $\mathbb{W}$ through ShortLex). So we have

$$\{0\} = \Theta_0 \leq \Theta_1 \leq \cdots \leq \Theta_{|\mathbb{W}|} = \mathrm{span}_K\{e_1, \dots, e_m\},$$

where the last equality is due to the fact that $M \supseteq \mathrm{EQ}_{\mathcal{B},w}$, the equations in $\mathrm{EQ}_{\mathcal{B},w}$ are supported on $\mathbb{W}$, the linear combinations of $e_1, \ldots, e_m$ given by the m equations in $\mathrm{EQ}_{\mathcal{B},w}$ are precisely the rows of $\mathcal{B}$, and $\mathcal{B}$ is regular by definition. Recall (Corollary 4.9) that there is a sole coincidence in $[e_i, e_i w]$ for every $i = 1, \ldots, m$, and let z_i denote the prefix of w so that $e_i z_i$ is the step in which the coincidence of $[e_i, e_i w]$ takes place.

Lemma 4.13. *We have $\Theta_{t-1} \lneq \Theta_t$ if and only if step t is a coincidence. In particular, if $g_i \in M|_{\mathbb{W}}$ is a (monic) element with leading monomial $e_i z_i$, then the vectors $\theta(g_1), \ldots, \theta(g_m)$ are linearly independent.*

Proof. We already explained why $\dim(\Theta_{|\mathbb{W}|}) = m$. Note that $\dim \Theta_t - \dim \Theta_{t-1} \in \{0, 1\}$, because every two monic elements $g_1, g_2 \in M|_{\mathbb{W}}$ with leading monomial v_t satisfy $\theta(g_1) - \theta(g_2) = \theta(g_1 - g_2) \in \Theta_{t-1}$. As there are exactly m coincidences, it is enough to prove that $\Theta_{t-1} = \Theta_t$ whenever step t is forced or free. If step t is free, then $M|_{D^{t-1}(\mathbb{W})} = M|_{D^t(\mathbb{W})}$ and obviously $\Theta_{t-1} = \Theta_t$. It thus remains to show that this is the case also if step t is forced.

Let ez be the monomial exposed in step t which is forced, and let $b \in B \cup B^{-1}$ be the edge leading to ez. There exists some $g \in M|_{D_b^t(\mathbb{W})}$ with ezb^{-1} in its support (in fact, its leading monomial), such that the coefficient of ezb^{-1} in g is $1 \in K$. If $\theta(g.b) \in \Theta_{t-1}$, then every other monic $f \in M|_{\mathbb{W}}$ with leading monomial ez satisfies $\theta(f) = \theta(f - g.b) + \theta(g.b) \in \Theta_{t-1}$ and we are done. So assume that $\theta(g.b) \notin \Theta_{t-1}$. In particular, $\theta(g.b) \neq 0$, so $g.b$ has some $e' \in E$ in its support, and so b^{-1} is the first letter of w. As w is assumed to be cyclically reduced, z is a proper prefix of w.

Now consider the monomial following ez in $[e, ew]$. Say it is ezc for some $b^{-1} \neq c \in B \cup B^{-1}$, and it is exposed at time s (so $s = t + m$). Because step t is forced, so is step s (by Corollary 4.9). As in the proof of Lemma 4.8, the monomials exposed between ez and ezc do not belong to $D_c^s(\mathbb{W})$, so $D_c^s(\mathbb{W}) \subseteq D^t(\mathbb{W})$. As step s is forced, there exists some monic $f \in M|_{D_c^s(\mathbb{W})} \subseteq M|_{D^t(\mathbb{W})}$ with ez its leading monomial. As before, as $\theta(f - g.b) \in \Theta_{t-1}$ but $\theta(g.b) \notin \Theta_{t-1}$, we get $\theta(f) = \theta(f - g.b) + \theta(g.b) \notin \Theta_{t-1}$. In particular, $\theta(f) \neq 0$. But $c \neq b^{-1}$ is not the first letter of w, so f cannot have any $e \in E$ in its support — a contradiction. This completes the proof of the first statement of the lemma. This also shows there exist $g_i \in M|_{\mathbb{W}}$ with leading monomial $e_i z_i$, for $i = 1, \ldots, m$, such that $\theta(g_1), \ldots, \theta(g_m)$ are linearly independent. The second statement of the lemma now follows from the fact that if $f, g \in M|_{\mathbb{W}}$ are both monic with leading monomial the t-th vertex, then $\theta(f) - \theta(g) \in \Theta_{t-1}$. □

Define $\mathbb{W}^2 \stackrel{\text{def}}{=} [e_1, e_1 w^2] \cup \cdots \cup [e_m, e_m w^2]$, and let $b \in B \cup B^{-1}$ be the first letter of w. For every $i = 1, \ldots, m$, let $f_{e_i z_i} \in M|_{\mathbb{W}}$ be the minimal monic element with leading monomial $e_i z_i$.

Lemma 4.14. *For every $i = 1, \ldots, m$, $f_{e_i z_i}$ is supported on $D_b(\mathbb{W}^2)$, and the outgoing b-edge at $e_i z_i$ is headed forward (towards $e_i w^2$).*

Proof. We proceed by induction on the order induced by ShortLex on $\{e_i z_i\}_{i=1,\ldots,m}$. The argument that follows works for both the base case and the induction step. If $z_i = w$, then there is an element in $\mathrm{EQ}_{\mathcal{B},w}$ with leading monomial $e_i w$ which is supported on $E \cup \{e_i w\}$, so $f_{e_i z_i}$ is also supported on $E \cup \{e_i w\}$ and the claim is clear. So assume that $|z_i| < |w|$, and that $e_i z_i$ is exposed at time t and admits

an outgoing c-edge towards $e_i w$. Then step $t+m$, in which $e_i z_i c$ is exposed, is forced, and there exists some $g \in M|_{D_c^{t+m}(\mathbb{W})} \subseteq M|_{D^t(\mathbb{W})}$ with leading monomial $e_i z_i$. By Lemma 4.13, $\theta(g) \notin \Theta_{t-1}$ so g has some $e \in E$ in its support, and therefore $c = b$.

Moreover, we may assume that g is supported on free steps and coincidences only. Indeed, the submodule M_{t+m-1} is generated on the free steps and coincidences exposed up to step $t+m-1$ (this is always the case in every valid exposure process), but by Corollary 4.9, in our case these vertices form a valid collection of subtrees $\mathbb{T}$ ($\mathbb{T} = T_1 \cup \cdots \cup T_m$, where $T_j = [e_j, e_j z_j] \cap [e_j, e_j z_i]$ for $j \geq i$ and $T_j = [e_j, e_j z_j] \cap [e_j, e_j z_i b]$ for $j < i$). But $e_i z_i b$ is forced, so every element with leading monomial $e_i z_i b$ belongs to M_{t+m-1}, and if we extend $\mathbb{T}$ to $e_i z_i b$ it is still a forced step (by Lemma 3.3). Thus there is some $g \in M|_{D_b(\mathbb{T} \cup \{e_i z_i b\})}$ with leading monomial $e_i z_i$.

If g has some coincidence $e_j z_j$ in its support other than $e_i z_i$, then as $e_j z_j < e_i z_i$, our induction hypothesis applies and $f_{e_j z_j} \in D_b(\mathbb{W}^2)$. Hence we may subtract $\alpha f_{e_j z_j}$ from g for some $\alpha \in K^*$ to decrease g, and $g - \alpha f_{e_j z_j} \in D_b(\mathbb{W}^2)$. If we repeat such subtractions as long as we can, we end up with a monic element f which is supported entirely on free vertices inside $D_b(\mathbb{W}^2)$ along with its leading monomial $e_i z_i$. Because all its nonleading monomials are free, this f is exactly $f_{e_i z_i}$ (otherwise $f - f_{e_i z_i} \neq 0$ is supported on free vertices, which is impossible), and we are done. □

Completing the proof of Theorem 1.12. Recall that $M \leq_{\mathbb{W}} \mathcal{A}^m$ is a fixed submodule satisfying $\operatorname{rk} M = m$ and $M \supseteq \mathrm{EQ}_{\mathcal{B},w}$. By Lemma 4.11, it is enough to show that M is generated by elements supported on $\{eu^j\}_{e \in E, j \in \{0,\ldots,d\}}$. By Theorem 3.8, $M = (f_{e_1 z_1}, \ldots, f_{e_m z_m})$, so it is enough to show that $f_{e_i z_i}$ is supported on $\{eu^j\}_{e \in E, j \in \mathbb{Z}}$ for all i.

Recall that b is the first letter of w. The submodule M contains $\mathrm{EQ}_{\mathcal{B},w}$ if and only if it contains $\mathrm{EQ}_{\mathcal{B},w}.b \stackrel{\text{def}}{=} \{f.b \mid f \in \mathrm{EQ}_{\mathcal{B},w}\}$. Define

$$\mathbb{W}^b \stackrel{\text{def}}{=} \bigcup_{e \in E} [b, wb],$$

and consider the exposure of M along $\mathbb{W}^b$ in the order induced from ShortLex. Clearly, the monomials that were free in the exposure along $\mathbb{W}$ are free now as well. We claim that the former coincidences $e_i z_i$ are now also free: as above, if $f \in M|_{\mathbb{W}^b}$ is monic with leading monomial $e_i z_i$, then $f_{e_i z_i} - f \in M$ is an element with $\theta(f_{e_i z_i} - f) = \theta(f_{e_i z_i})$ but with leading monomial smaller than $e_i z_i$, which contradicts Lemma 4.13. On the other hand, by Lemma 4.14, $f_{e_i z_i}.b \in M|_{\mathbb{W}^b}$ has leading monomial $e_i z_i b$, and so $e_i z_i b$ is a coincidence in the exposure of M along $\mathbb{W}^b$. Moreover, the nonleading monomials of $f_{e_i z_i}.b$ are all free in the exposure along $\mathbb{W}^b$, so $f_{e_i z_i}.b$ is the minimal monic element in $M|_{\mathbb{W}^b}$ with leading monomial $e_i z_i b$. The same argument as in Lemma 4.14 shows that $f_{e_i z_i}.b$ is supported on $D_c(\mathbb{W}^2)$ and the outgoing c-edge at $e_i z_i b$ is headed towards $e_i w^2$, where $c \in B \cup B^{-1}$ is the second letter of w.

This argument can now go on to the exposure of M along $\mathbb{W}^{bc}$ and so on, and shows that for every prefix w' of w and every i, $f_{e_i z_i}.w'$ is supported on $[1, w^2]$. This completes the proof exactly as in the proof of Theorem 1.2 in Section 4A. □

5. The quotient module $K[F]/(w-1)$

Fix $w \in \boldsymbol{F}$, and consider the right $\mathcal{A}$-module obtained as a quotient of the $\mathcal{A}$-module $\mathcal{A}$ by its submodule $(w-1)$. We denote this quotient by

$$\mathcal{A}_w \stackrel{\text{def}}{=} K[\boldsymbol{F}]/(w-1) = \mathcal{A}/(w-1).$$

We study this module and prove two main results about it. First, we show that if w is a nonpower, then the only cyclic generators of $\mathcal{A}_w$ are the "obvious ones" (Theorem 5.4). Second, we prove that whenever a subtree $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ supports both $w-1$ and a rank-2 ideal $I \leq_T \mathcal{A}$ in which $w-1$ is primitive, there is an element $f \in \mathcal{A}$ supported on T so that $\{f, w-1\}$ is a basis of I (Corollary 5.2). In particular, the latter result yields an algorithm to test whether $w-1$ is primitive in a given rank-2 ideal (Corollary 5.3). We need these two results for our proof of Theorem 1.4 in Section 6, but we also find them interesting for their own right. See Section 7 for a discussion on potential generalizations of these results.

Consider the Schreier graph

$$\mathcal{S}_w \stackrel{\text{def}}{=} \mathrm{Sch}(\boldsymbol{F} \curvearrowright \langle w \rangle \backslash \boldsymbol{F}, B) = \langle w \rangle \backslash \mathrm{Cay}(\boldsymbol{F}, B).$$

This is a graph whose vertices correspond to the right cosets of the subgroup $\langle w \rangle$ in $\boldsymbol{F}$. For every vertex $\langle w \rangle z$ and every $b \in B$, there is a directed b-edge from the vertex $\langle w \rangle z$ to the vertex $\langle w \rangle zb$. In other words, this is the quotient of $\mathrm{Cay}(\boldsymbol{F}, B)$ by the action of $\langle w \rangle$ from the left. Note that $\mathcal{S}_w$ is made of a cycle (reading the cyclic reduction of w) with infinite trees hanging from it (unless $\mathrm{rk}\, \boldsymbol{F} = 1$, in which case $\mathcal{S}_w$ is a mere cycle). This is illustrated in Figure 1.

An element $f \in \mathcal{A}$ belongs to the ideal $(w-1)$ if and only if for every $z \in \boldsymbol{F}$, the coefficients in f of the elements in the right coset $\langle w \rangle z$ sum up to zero. Therefore, the elements of $\mathcal{A}_w$ are given by K-linear combinations of right cosets of $\langle w \rangle$, namely, K-linear combinations of the vertices of $\mathcal{S}_w$. This can also

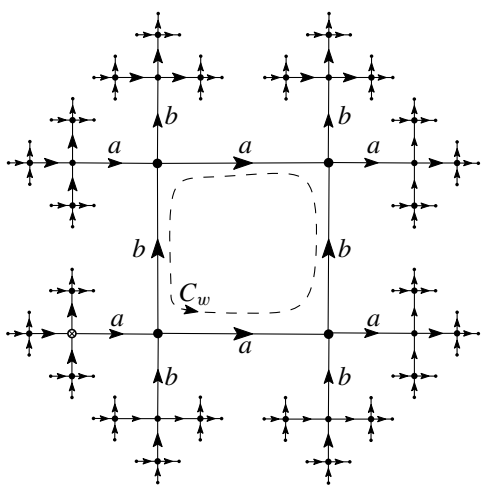


Figure 1. The Schreier graph $\mathcal{S}_w$ for $w = a[a, b]a^{-1}$. The unique simple cycle is marked by C_w.

be seen by the fact that a possible Schreier transversal of the ideal $(w-1)$ is obtained by considering $\mathrm{Cay}(\boldsymbol{F}, B)$, cutting the axis[10] of w on both sides of one period of the cyclic reduction of w, and taking the connected component of this period.

Now consider the quotient map

$$\rho : \mathcal{A} \twoheadrightarrow \mathcal{A}_w,$$

which, by abuse of notation, we also regard as the graph morphism

$$\rho : \mathrm{Cay}(\boldsymbol{F}, B) \to \mathcal{S}_w.$$

Whenever a subtree $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ contains $[1, w]$, its image $\rho(T) \subseteq \mathcal{S}_w$ contains the cycle in $\mathcal{S}_w$. In fact, it suffices that T contains any interval in the axis of w of length at least the length of the cyclic reduction of w.

Lemma 5.1. *Let $G \subseteq \mathcal{S}_w$ be a connected subgraph which contains the cycle of $\mathcal{S}_w$. Let $f \in \mathcal{A}_w$ satisfy that none of $\{f.z \mid z \in \boldsymbol{F}\}$ is supported on G. Then the submodule $f\mathcal{A} \leq \mathcal{A}_w$ does not contain any nonzero element supported on G.*

Proof. On the vertices of $\mathcal{S}_w \setminus G$ define an "exploration" as in Definition 3.2: this is an enumeration of these vertices such that every vertex is a neighbor of some vertex in G or of a smaller vertex. This exploration induces a preorder on the orbit $\{f.z \mid z \in \boldsymbol{F}\}$ obtained by comparing the largest vertex in their support with respect to this exploration order (by assumption, every element $f.z$ in this orbit has at least one vertex outside G in its support). Assume without loss of generality that f is an element of the orbit with the smallest possible maximal vertex in its support. Denote this vertex $v_{\max}$. Denote by $\overline{G}$ the (connected) subgraph of $\mathcal{S}_w$ consisting of G together with the prefix $\{v \in \mathrm{vert}(\mathcal{S}_w \setminus G) \mid v \leq v_{\max}\}$ of the exploration order on $\mathcal{S}_w \setminus G$.

Now consider the element $fg \in \mathcal{A}_w$ for an arbitrary $g \in \mathcal{A}$ not supported on the identity $e \in \boldsymbol{F}$. It suffices to show that fg is not supported on $\overline{G}$ (let alone on G). Write $f = \alpha_1 f_1 + \cdots + \alpha_m f_m$ with $\alpha_1, \ldots, \alpha_m \in K^*$ and distinct $f_1, \ldots, f_m \in \mathrm{vert}(\mathcal{S}_w)$, and write $g = \beta_1 g_1 + \cdots + \beta_\ell g_\ell$ with $\beta_1, \ldots, \beta_\ell \in K^*$ and distinct $g_1, \ldots, g_\ell \in \boldsymbol{F}$ and so that $|g_1| \geq |g_2| \geq \cdots \geq |g_\ell|$. Denote by $b \in B \cup B^{-1}$ the first letter in g_1. Then $f.b$ cannot be supported on $\overline{G}$: otherwise, $f.b$ would be supported on G together with vertices strictly smaller than $v_{\max}$ in $\mathcal{S}_w \setminus G$ (we use here the fact that $v_{\max}$ is a leaf in $\overline{G}$), contradicting our assumption about f. So there is a monomial f_i in the support of f such that $f_i.b$ is a monomial outside $\overline{G}$. But then $f_i g_1$ is at distance $|g_1|$ from $\overline{G}$, with the closest vertex of $\overline{G}$ being f_i. Clearly, $f_i g_1 \neq f_j g_k$ for every $(j,k) \neq (i,1)$, because the only path of length $|g_1|$ from $\overline{G}$ to $f_i g_1$ in $\mathcal{S}_w$, is the path starting at f_i and reading g_1. Thus $f_i g_1$ belongs to the support of fg, and fg is not supported on $\overline{G}$. □

Corollary 5.2. *Let $1 \neq w \in \boldsymbol{F}$ and $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ be a subtree which contains $[1, w]$. Assume that $I \leq_T \mathcal{A}$ is a rank-2 ideal supported on T which contains $w-1$ as a primitive element. Then there is an element $f \in \mathcal{A}$ supported on T so that $\{f, w-1\}$ is a basis for I.*

[10] The axis of w is composed of the points in $\mathrm{Cay}(\boldsymbol{F}, B)$ moved by left multiplication by w by the least distance.

Proof. As $w-1$ is primitive in I, there is some $f \in \mathcal{A}$ which completes it to a basis of I. Consider $\overline{T} = \rho(T)$, the image of T in $\mathcal{S}_w$ and let $\bar{f} = \rho(f) \in \mathcal{A}_w$. If $\{f, w-1\}$ is a basis for I, then so is $\{g, w-1\}$ for every $g \in \rho^{-1}(\bar{f})$, because in this case $f - g \in (w-1)$. So if $\bar{f}.z$ is supported on $\overline{T}$ for some $z \in \boldsymbol{F}$, we are done: if $\{f, w-1\}$ is a basis then so is $\{f.z, w-1\}$. Otherwise, we are in the situation of Lemma 5.1, and $\bar{f}\mathcal{A}$ does not contain any element supported on $\overline{T}$. But $\bar{f}\mathcal{A}$ contains $\rho(I)$ (in fact $\bar{f}\mathcal{A} = \rho(I)$), and as I is generated on T, I contains an element $h \in I \setminus (w-1)$ which is supported on T. Then $\bar{f}\mathcal{A} \ni \rho(h)$, which is a contradiction as $\rho(h) \neq 0$ and is supported on $\overline{T}$. □

Corollary 5.3. *If the field K is finite,*[11] *there is an algorithm to test, given a (generating set of a) rank-2 ideal $I \leq \mathcal{A}^m$ and a word $w \in \boldsymbol{F}$, whether $w-1$ is primitive in I.*

Proof. By [Cohn 1964, Proposition 2.2], every pair of generators of I is a basis. So $\{f, w-1\}$ is a basis for I if and only if $f, w-1 \in I$ and $(f, w-1)$ contains the given generating set of I. By Corollary 5.2, $w-1$ is primitive in I if and only if there exists an element f supported on T such that $\{f, w-1\}$ is a basis of I. As K is finite, there are finitely many elements supported on T. Finally, Rosenmann [1993] describes an algorithm to test whether a given element belongs to a given ideal in $\mathcal{A}$ (where the ideal is given by a finite generating set). □

Corollaries 5.2 and 5.3 naturally raise the question to what extent they can be generalized for ideals of rank larger than two and for elements of $\mathcal{A}$ which are not of the form $w-1$ — see Section 7 for a discussion around it.

5A. *Cyclic generators of $K[F]/(w-1)$.* The group algebra $\mathcal{A} = K[\boldsymbol{F}]$ has only trivial units — a scalar times an element of the group[12] (this property was conjectured by Kaplansky to hold in all group algebras of torsion-free groups over fields but a counterexample has recently been found [Gardam 2021]). The goal of this subsection is to prove a similar result for $\mathcal{A}_w = \mathcal{A}/(w-1)$. While $\mathcal{A}_w$ is not a ring and therefore does not admit units, it does admit cyclic generators as an $\mathcal{A}$-module: elements $f \in \mathcal{A}_w$ such that $f\mathcal{A} = \mathcal{A}_w$. Clearly, for every unit of $\mathcal{A}$, its image in $\mathcal{A}_w$ is a cyclic generator. Here we prove that provided that w is not a power, all cyclic generators of $\mathcal{A}_w$ are of this sort.

Theorem 5.4. *Assume that $1 \neq w \in \boldsymbol{F}$ is a nonpower. Then every cyclic generator of the right $\mathcal{A}$-module $\mathcal{A}_w = \mathcal{A}/(w-1)$ is an image of a unit of $\mathcal{A}$.*

Namely, every cyclic generator of $\mathcal{A}_w$ is a coset of the form $\alpha z + (w-1)$ for some $\alpha \in K^*$ and $z \in \boldsymbol{F}$.

Remark 5.5. Theorem 5.4 is false for proper powers. For example, if $|K| = 3$ and $w = a^3$, then $\rho(a+1) \in \mathcal{A}_w$ is not a ρ-image of a unit of $\mathcal{A}$: its support in $\mathcal{S}_w$ is of size two. Yet $a^3 + 1 \in (a+1)$ and so $\rho(2) = \rho(a^3+1) \in \rho(a+1)\mathcal{A}$. Thus $\rho(a+1)$ is a cyclic generator of $\mathcal{A}_w$.

[11] We assume throughout the paper that K is finite, but some of the results about free group algebras, such as Corollary 5.2, hold for infinite fields just as well. In contrast, Corollary 5.3 relies on K being finite.

[12] This is well known. It can also be seen, for example, by an argument similar to the one in the proof of Lemma 5.1: for any $0 \neq f \in \mathcal{A}$ with support of size at least 2, take a minimal subtree T of $\mathcal{C} = \mathrm{Cay}(\boldsymbol{F}, B)$ which supports an element in the orbit $\{f.z \mid z \in \boldsymbol{F}\}$. Then the argument in the proof of Lemma 5.1 shows that $f\mathcal{A}$ does not contain elements supported on T except for scalar multiples of f.

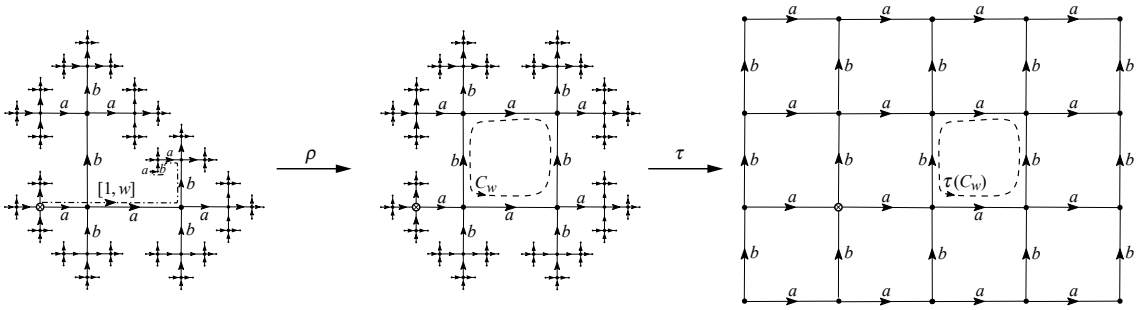


Figure 2. Let $w = a[a, b]a^{-1}$. The Cayley graph of $\boldsymbol{F} = \boldsymbol{F}(a, b)$ is on the left with $[1, w]$ marked. The middle graph is $\mathcal{S}_w$, and the graph at the right side is a piece of the Cayley graph of $\boldsymbol{F}/\langle\langle w\rangle\rangle$. In all graphs, the vertex corresponding to the identity element or its ρ-image is marked with $\otimes$.

First we show that cyclic generators in $\mathcal{A}_w$ may be assumed to be supported on the cycle of $\mathcal{S}_w$.

Lemma 5.6. *If $f \in \mathcal{A}_w$ is a cyclic generator of $\mathcal{A}_w$, then there is some $z \in \boldsymbol{F}$ such that fz is supported on the cycle in $\mathcal{S}_w$.*

Proof. This follows immediately from Lemma 5.1 applied to G being the cycle in $\mathcal{S}_w$. □

Let $w \in \boldsymbol{F}$ be a nonpower. If w' is the cyclic reduction of w then the automorphism of $\boldsymbol{F}$ mapping w to w' extends to an automorphism of $\mathcal{A}$ and induces isomorphisms $\mathcal{A}_w \xrightarrow{\cong} \mathcal{A}_{w'}$ and $\mathcal{S}_w \xrightarrow{\cong} \mathcal{S}_{w'}$. Thus we may assume without loss of generality that w is cyclically reduced. Denote by $\langle\langle w\rangle\rangle$ the normal closure of w in $\boldsymbol{F}$, and denote by $((w-1))$ the two-sided ideal of $\mathcal{A}$ generated by $w-1$. Since we have $\{0\} \subseteq (w-1) \subseteq ((w-1))$, we get canonical epimorphisms of right $\mathcal{A}$-modules

$$\mathcal{A} \xrightarrow{\rho} \!\!\!\rightarrow \mathcal{A}_w \xrightarrow{\tau} \!\!\!\rightarrow \mathcal{A}/((w-1)).$$

See Figure 2.

Lemma 5.7. *Let $p : \boldsymbol{F} \to \boldsymbol{F}/\langle\langle w\rangle\rangle$ be the canonical projection. The map $\varphi : \mathcal{A}/((w-1)) \to K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$ defined by $\sum_{z\in F} \alpha_z z + ((w-1)) \mapsto \sum_{z\in F} \alpha_z p(z)$ is an isomorphism of K-algebras.*

Proof. The proof is a standard argument in algebra, but we include it for completeness. By the universal property of group rings, the group homomorphism $p : \boldsymbol{F} \to \boldsymbol{F}/\langle\langle w\rangle\rangle \subseteq K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$ extends to a unique K-algebra epimorphism $\psi : \mathcal{A} \twoheadrightarrow K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$. The ideal $((w-1))$ lies in the kernel of ψ: it is enough to show that $u(w-1)v \in \ker\psi$ for $u, v \in \boldsymbol{F}$, and

$$\psi(u(w-1)v) = \psi(uwv - uv) = \psi(uwv) - \psi(uv) = p(uwv) - p(uv) = 0,$$

where the last equality is because uwv and uv lie in the same coset of $\langle\langle w\rangle\rangle$. Thus, the homomorphism ψ induces an epimorphism $\psi' : \mathcal{A}/((w-1)) \twoheadrightarrow K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$. For every $z \in \boldsymbol{F}$, ψ' satisfies $\psi'(z+((w-1))) = \psi(z) = p(z)$, and so by linear extension ψ' agrees with φ from the statement of the lemma (in particular, φ is a well-defined epimorphism of K-algebras).

It is left to show that φ is injective. Suppose that

$$\varphi\Big(\sum_{z\in F}\alpha_z z + ((w-1))\Big) = \sum_{z\in F}\alpha_z p(z) = 0.$$

For every coset C of $\langle\langle w\rangle\rangle$ we have $\sum_{z\in C}\alpha_z = 0$. We complete the proof by showing that this implies that $\sum_{z\in C}\alpha_z z \in ((w-1))$ — and then the sum over all cosets would also lie in $((w-1))$. Such a finite sum can always be decomposed as a sum over elements of the form $\alpha(z_2 - z_1)$ where $\alpha \in K$ and $z_1, z_2 \in C$. In every such element, z_2 can be obtained from z_1 by a finite sequence of multiplications from the right by conjugates of w or w^{-1}, and so it is enough to show that $z_2 - z_1 \in ((w-1))$ for $z_2 = z_1 \cdot uw^{\varepsilon}u^{-1}$ where $u \in F$ and $\varepsilon \in \{\pm 1\}$. And indeed we have $z_2 - z_1 = z_1 u(w^{\varepsilon} - 1)u^{-1} \in ((w-1))$. □

In our proof of Theorem 5.4 we use the following well-known concept.

Definition 5.8. A *right order* on a group Γ is a linear order on Γ such that for every $r, s, t \in \Gamma$ with $r < s$ we have $rt < st$. A group is called *right-orderable* if it admits a right order.

It is well-known that Kaplansky's unit conjecture, mentioned above, is true for right orderable groups — see, e.g., [Clay and Rolfsen 2016, Theorem 1.58]. We add a proof for completeness.

Lemma 5.9. *Let K be a field and Γ a right-orderable group. If $ts = 1$ for $t, s \in K[\Gamma]$ then $t = \lambda g$ for some $\lambda \in K^*$ and $g \in \Gamma$.*

Proof. Write $t = \sum_{i=1}^{n}\lambda_i g_i$ for $\lambda_i \in K^*$ and $g_1, \dots, g_n \in \Gamma$ distinct. Since $ts = 1$ we know that $n \neq 0$. Now assume towards contradiction that $n \geq 2$. Let $<$ be a right order for Γ. Assume without loss of generality that $g_1 < g_2 < \cdots < g_n$. Write similarly $s = \sum_{j=1}^{m}\mu_j h_j$ for $h_1 < h_2 < \cdots < h_m$ and $\mu_j \in K^*$. Then we have $1 = ts = \sum_{i,j}\lambda_i\mu_j g_i h_j$.

We now find two elements of Γ such that their coefficients in ts are nonzero. Let $j_{\min}$ be the index such that $g_1 h_{j_{\min}} = \min\{g_1h_1, g_1h_2, \dots, g_1h_m\}$. In particular, $g_1 h_{j_{\min}}$ is strictly smaller than any other $g_1 h_j$ for $j \neq j_{\min}$. In addition, if $i \neq 1$ then $g_1 h_{j_{\min}} \leq g_1 h_j < g_i h_j$. Thus, the coefficient of $g_1 h_{j_{\min}}$ in ts is $\lambda_1\mu_{j_{\min}} \neq 0$. Similarly, let $j_{\max}$ be the index such that $g_n h_{j_{\max}} = \max\{g_nh_1, g_nh_2, \dots, g_nh_m\}$. A similar argument shows that the coefficient of $g_n h_{j_{\max}}$ in ts is $\lambda_n\mu_{j_{\max}} \neq 0$. Finally, since $n \geq 2$, we have $g_1 h_{j_{\min}} < g_n h_{j_{\min}} \leq g_n h_{j_{\max}}$ and so $g_1 h_{j_{\min}}$ and $g_n h_{j_{\max}}$ are distinct elements of Γ with nonzero coefficients in $ts = 1$ — a contradiction. □

The following theorem is a well-known result in the theory of one-relator groups.

Theorem 5.10. *If $1 \neq w \in \boldsymbol{F}$ is a nonpower then the one-relator group $\boldsymbol{F}/\langle\langle w\rangle\rangle$ is right-orderable.*

Proof. As w is a nonpower, we deduce that $\boldsymbol{F}/\langle\langle w\rangle\rangle$ is torsion-free by a theorem of Karass, Magnus and Solitar [Karrass et al. 1960, Theorem 1]. By a theorem proven independently by Brodskii [1984, Corollary 2.3] and Howie [1982, Corollary 4.3], every torsion-free one-relator group has the property of being *locally indicable*, which means that each of its nontrivial finitely generated subgroups admits a nontrivial homomorphism to $\mathbb{Z}$. Finally, the Burns–Hale theorem [1972, Theorem 2] states that a group H is right-orderable if and only if any nontrivial finitely generated subgroup of H admits a nontrivial

homomorphism to some right-orderable group. Since $\mathbb{Z}$ is right-orderable (the usual order on $\mathbb{Z}$ is a right order), the Burns–Hale theorem implies that every locally indicable group is right-orderable. The combination of the theorems above gives that $\boldsymbol{F}/\langle\langle w\rangle\rangle$ is right-orderable. □

Proof of Theorem 5.4. Let $1 \neq w \in \boldsymbol{F}$ be a nonpower and suppose that $\bar{f} \in \mathcal{A}_w$ generates $\mathcal{A}_w$. Since ρ is surjective, there exists some $f \in \mathcal{A}$ such that $\rho(f) = \bar{f}$. As $\bar{f}$ generates $\mathcal{A}_w$, there exists some $s \in \mathcal{A}$ such that $\bar{f}s = \rho(1)$ or, equivalently, $\rho(fs) = \rho(1)$. Applying τ to both sides of the equation and using the fact that $\tau \circ \rho$ is a homomorphism of K-algebras we obtain $\tau\rho(f) \cdot \tau\rho(s) = \tau\rho(1)$, and in particular $\tau\rho(f)$ has a right inverse in the quotient K-algebra $\mathcal{A}/((w-1))$. Now, since $\tau\rho(f)$ has a right inverse in $\mathcal{A}/((w-1))$, its image under the isomorphism φ from Lemma 5.7 has a right inverse in $K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$. By Theorem 5.10, as w is not a power, $\boldsymbol{F}/\langle\langle w\rangle\rangle$ is right-orderable. Lemma 5.9 applied for $\Gamma = \boldsymbol{F}/\langle\langle w\rangle\rangle$, implies that $\varphi(\tau\rho(f)) = \lambda g$ for some $\lambda \in K^*$ and $g \in \boldsymbol{F}/\langle\langle w\rangle\rangle$.

Without loss of generality, by Lemma 5.6, we may assume that $\bar{f} = \rho(f)$ is supported on cosets of $\langle w\rangle$ belonging to the unique simple cycle of $\mathcal{S}_w$. The Weinbaum subword theorem [1972, Theorem 2] asserts that none of the nontrivial proper subwords of the cyclic reduction of w lies in its normal closure $\langle\langle w\rangle\rangle$. This implies that two distinct vertices of the cycle of $\mathcal{S}_w$ have distinct images through τ, namely, their images belong to different elements of $\boldsymbol{F}/\langle\langle w\rangle\rangle$. But the τ-image of $\bar{f}$ is λg, which is supported on a single element $g \in K[\boldsymbol{F}/\langle\langle w\rangle\rangle]$. Thus $\bar{f}$ itself is supported on a single element of the cycle of $\mathcal{S}_w$ and can be lifted to an element $f \in \mathcal{A}$ supported on a single element of $\boldsymbol{F}$. □

6. Critical ideals of rank 2

Throughout this section fix a nonpower $1 \neq w \in \boldsymbol{F}$ and assume without loss of generality that it is cyclically reduced. Theorems 1.1 and 1.2 yield that $\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{c}{q^N} + O\left(\frac{1}{q^{2N}}\right)$ for some constant c. Our goal is to prove Theorem 1.4:

$$c = |\mathrm{Crit}_q^2(w)|,$$

where $\mathrm{Crit}_q^2(w)$ is the set of rank-2 ideals $I \leq \mathcal{A}$ containing $w-1$ as an imprimitive element.

Recall our formula (2-5) for $\mathbb{E}_w[\mathrm{fix}]$ and Corollary 3.9. It follows that the $\frac{1}{q^N}$-coefficient of $\mathbb{E}_w[\mathrm{fix}]$ consists of the contributions of the rank-1 and rank-2 ideals in the set

$$\mathcal{I} \stackrel{\mathrm{def}}{=} \{I \leq_{[1,w]} \mathcal{A} \mid I \ni w-1\}.$$

As $w \neq 1$ and is a nonpower, by Corollary 4.6 the rank-1 ideals in $\mathcal{I}$ are precisely (1) and $(w-1)$. The contribution of (1) to (2-5) is precisely 1, so it does not affect c. Denote by β_w the coefficient of $\frac{1}{q^N}$ in the contribution of $(w-1)$, namely, this contribution is $1 + \frac{\beta_w}{q^N} + O\left(\frac{1}{q^{2N}}\right)$. The summand in (2-5) corresponding to a rank-2 ideal is $\frac{1}{q^N} + O\left(\frac{1}{q^{2N}}\right)$, so such an ideal contributes exactly 1 to c. Recall that all the ideals in $\mathrm{Crit}_q^2(w)$ are in $\mathcal{I}$, by Corollary 3.11. Denote by $\mathrm{Prim}^2(w)$ the set of rank-2 ideals in $\mathcal{I}$ in which $w-1$ is primitive. With this notation, the coefficient c of $\frac{1}{q^N}$ in $\mathbb{E}_w[\mathrm{fix}]$ is $c = \beta_w + |\mathrm{Prim}^2(w)| + |\mathrm{Crit}_q^2(w)|$. Our goal is, thus, to prove that

$$\beta_w + |\mathrm{Prim}^2(w)| = 0.$$

Recall from Section 5 the quotient $\mathcal{A}$-module $\mathcal{A}_w \stackrel{\text{def}}{=} \mathcal{A}/(w-1)$, the projection $\rho : \mathcal{A} \to \mathcal{A}_w$ and the Schreier graph $\mathcal{S}_w = \langle w \rangle \backslash \mathrm{Cay}(\boldsymbol{F}, B)$. The elements of $\mathcal{A}_w$ are K-linear combinations of the vertices of $\mathcal{S}_w$, and we use ρ to denote also the quotient in the graph level $\rho : \mathrm{Cay}(\boldsymbol{F}, B) \to \mathcal{S}_w$. Let $C_w = \rho([1, w])$ denote the unique simple cycle in $\mathcal{S}_w$ (here we use the fact that w is assumed to be cyclically reduced).

Lemma 6.1. *The $\frac{1}{q^N}$-coefficient of the summand corresponding to $I = (w-1)$ in (2-5) is*

$$\beta_w = -\frac{q^{v(C_w)} - 1}{q-1} + \sum_{b \in B} \frac{q^{e_b(C_w)} - 1}{q-1}. \tag{6-1}$$

Proof. Recall that β_w is the $\frac{1}{q^N}$-coefficient of the Laurent expansion of

$$\frac{\mathfrak{indep}_{|w|+1-d^{[1,w]}(I)}(V_N)}{\prod_{b \in B} \mathfrak{indep}_{e_b(w)-d_b^{[1,w]}(I)}(V_N)}, \tag{6-2}$$

for $I = (w-1)$. Because $f \in \mathcal{A}|_{[1,w]}$ belongs to $I = (w-1)$ if and only if its coefficients in every fiber over C_w sum up to zero, the dimension over K of $I|_{[1,w]}$ is precisely $d^{[1,w]}(I) = v([1,w]) - v(C_w) = 1$. Similarly, the dimension over K of $I|_{D_b([1,w])}$ is precisely $d_b^{[1,w]}(I) = e_b([1,w]) - e_b(C_w) = 0$. Hence, (6-2) is equal to

$$\frac{\mathfrak{indep}_{v(C_w)}(V_N)}{\prod_{b \in B} \mathfrak{indep}_{e_b(C_w)}(V_N)} = \frac{(q^N-1)(q^N-q)\cdots(q^N-q^{v(C_w)-1})}{\prod_{b \in B}(q^N-1)(q^N-q)\cdots(q^N-q^{e_b(C_w)-1})}. \tag{6-3}$$

Because C_w is a cycle, the number of vertices is identical to the total number of edges. Hence (6-3) is equal to

$$\frac{\left(1-\frac{1}{q^N}\right)\left(1-\frac{q}{q^N}\right)\cdots\left(1-\frac{q^{v(C_w)-1}}{q^N}\right)}{\prod_{b \in B}\left(1-\frac{1}{q^N}\right)\left(1-\frac{q}{q^N}\right)\cdots\left(1-\frac{q^{e_b(C_w)-1}}{q^N}\right)},$$

and the $\frac{1}{q^N}$-coefficient of the Laurent expansion of this expression is

$$(-1-q-\cdots-q^{v(C_w)-1}) - \sum_{b \in B}(-1-q-\cdots-q^{e_b(C_w)-1}),$$

which is equal to (6-1). □

Denote by D_w the set of proper nontrivial cyclic submodules[13] of $\mathcal{A}_w$ generated by some element supported on the cycle C_w:

$$D_w \stackrel{\text{def}}{=} \{g\mathcal{A} \lneq \mathcal{A}_w \mid 0 \neq g \in \mathcal{A}_w \text{ and } g \text{ supported on } C_w\}.$$

Lemma 6.2. *There is a one-to-one correspondence between D_w and* $\mathrm{Prim}^2(w)$.

[13] Recall that a cyclic submodule is a submodule generated by a single element.

Proof. By Corollary 5.2, the rank-2 ideals $I \leq_{[1,w]} \mathcal{A}$ containing $w-1$ as a primitive element are exactly the rank-2 ideals of the form $(w-1, f)$ with f supported on $[1, w]$ (here we use again the fact that every pair of generators of a rank-2 ideal is a basis — [Cohn 1964, Proposition 2.2]). Now $g \in \mathcal{A}_w$ is supported on C_w if and only if there is some $f \in \mathcal{A}$ supported on $[1, w]$ with $\rho(f) = g$. Note that

$$(w-1, f) = \rho^{-1}(\rho(f)\mathcal{A}),$$

where $\rho(f)\mathcal{A}$ is the submodule of $\mathcal{A}_w$ generated by the image of f in A_w. Because the only rank-1 ideals containing $w-1$ are (1) and $(w-1)$, we have that $(w-1, f)$ is of rank 2 if and only if $\rho(f)\mathcal{A}$ is a nonzero proper submodule of $\mathcal{A}_w$. □

Next, we study the different elements $g \in \mathcal{A}_w$ supported on C_w. In order to understand when two different elements g, g' generate the same submodule, we construct a graph Υ. The vertices of Υ are the 1-dimensional linear subspaces of $K^{\mathrm{vert}(C_w)}$, so their number is $v(\Upsilon) = \frac{q^{v(C_w)}-1}{q-1}$. For every $b \in B$ and every 1-dimensional subspace $U \leq K^{b\text{-edges}(C_w)}$ (here $K^{b\text{-edges}}$ is the space of K-linear combinations of the b-edges in C_w), the subspace U corresponds to a 1-dimensional subspace $o(U)$ of the vertices supported on the origins of the b-edges, as well as a 1-dimensional subspace $t(U)$ supported on the termini of the b-edges. For every such U we draw a directed b-edge from the vertex $o(U)$ to the vertex $t(U)$ in Υ. Note that $e(\Upsilon) = \sum_{b \in B} \frac{q^{e_b(C_w)}-1}{q-1}$, so overall

$$\chi(\Upsilon) \stackrel{\mathrm{def}}{=} v(\Upsilon) - e(\Upsilon) = -\beta_w, \tag{6-4}$$

where the second equality is by Lemma 6.1. Denote by $\mathcal{C}(\Upsilon)$ the connected components of Υ. Because $g\mathcal{A} = g.b\mathcal{A}$ for every $g \in \mathcal{A}_w$ and $b \in B$, there is a well-defined surjective map

$$\Phi : \mathcal{C}(\Upsilon) \twoheadrightarrow \{g\mathcal{A} \mid 0 \neq g \in \mathcal{A}_w \text{ supported on } C_w\} = D_w \cup \{\mathcal{A}_w\}.$$

One of the connected components of Υ is isomorphic to C_w: this is the component consisting of vertices and edges of Υ corresponding to 1-dimensional subspaces supported on a single vertex or on a single edge. Denote this component by C_0. Clearly, $\Phi(C_0) = \mathcal{A}_w$.

Lemma 6.3. *All connected components of Υ except for C_0 are paths.*

Proof. The degree of every vertex in Υ is at most 2, so every connected component is a path or a cycle. Assume that some component $C_0 \neq C \in \mathcal{C}(\Upsilon)$ is a cycle. Let U be a vertex in C, and assume that this cycle reads the (cyclically reduced) word $z \in \boldsymbol{F}$ starting (and ending) at U. Recall that U is a 1-dimensional subspace of $K^{\mathrm{vert}(C_w)}$ supported on at least two vertices of C_w, and denote the support of U by $\mathrm{supp}(U)$, so $|\mathrm{supp}(U)| \geq 2$. In particular, for every $s \in \mathrm{supp}(U)$, there is a path in C_w reading z leaving s and reaching some $s' \in \mathrm{supp}(U)$. Hence, some power z^k of z is a path from s to itself for every $s \in \mathrm{supp}(U)$. Because w is not conjugate to w^{-1}, every such copy of z^k has the same orientation along C_w. We get that there is some $y \in \boldsymbol{F} \setminus \langle w \rangle$ so that $ywy^{-1} = w$. This is not possible unless w is a proper power, which is not the case. □

Lemma 6.4. *The map $\Phi : \mathcal{C}(\Upsilon) \to D_w \cup \{\mathcal{A}_w\}$ is one-to-one.*

Proof. Theorem 5.4 states the only cyclic generators of $\mathcal{A}_w$ are elements supported on a single vertex of $\mathcal{S}_w$, and so C_0 is the only connected component in Υ mapped to $\mathcal{A}_w$. It remains to show that every element of D_w has a single preimage in $\mathcal{C}(\Upsilon)$. Suppose that $g, g' \in \mathcal{A}_w$, both supported on C_w, so that $g'\mathcal{A} = g\mathcal{A} \in D_w$, namely, $\{0\} \neq g\mathcal{A} = g'\mathcal{A} \lneqq \mathcal{A}_w$. Let $f, f' \in \mathcal{A}$ be preimages of g, g', respectively, through ρ^{-1}, which are supported on $[1, w]$. Then $(f, w-1) = (f', w-1)$ is a rank-2 ideal by Lemma 6.2. Thus there are $p_1, p_2, q_1, q_2 \in \mathcal{A}$ such that

$$f' = fp_1 + (w-1)p_2, \quad f = f'q_1 + (w-1)q_2,$$

so

$$f = (fp_1 + (w-1)p_2)q_1 + (w-1)q_2 = f \cdot p_1q_1 + (w-1)(p_2q_1 + q_2).$$

But $\{f, w-1\}$ is a basis, so by uniqueness we get $p_1q_1 = 1$ (and $p_2q_1 + q_2 = 0$). The only units of $\mathcal{A}$ are scalar product of monomials of the form αz with $\alpha \in K^*$ and $z \in \boldsymbol{F}$ (this was mentioned and explained in Section 5A). By multiplying g' by a scalar if necessary, we may thus assume that $p_1 \in \boldsymbol{F}$ is a word, and we get that

$$g' = \rho(f') = \rho(fp_1 + (w-1)p_2) = \rho(fp_1) = \rho(f)p_1 = gp_1.$$

But C_w contains every reduced path between every two of its vertices, so inside Υ there is a path (reading p_1) from the vertex corresponding to g to the one corresponding to g'. In particular, they both belong to the same connected component. □

Completing the proof of Theorem 1.4. Recall that we need to show that $\beta_w + |\mathrm{Prim}^2(w)| = 0$. Consider the above mentioned map $\Phi : \mathcal{C}(\Upsilon) \to D_w \cup \{\mathcal{A}_w\}$. As C_0 is a cycle isomorphic to C_w, we have $\chi(C_0) = 0$. By Lemma 6.3, $\chi(C) = 1$ for any $C_0 \neq C \in \mathcal{C}(\Upsilon)$, so $|\mathcal{C}(\Upsilon) \setminus \{C_0\}| = \chi(\Upsilon)$. Thus

$$|\mathrm{Prim}^2(w)| = |D_w| = |\mathcal{C}(\Upsilon) \setminus \{C_0\}| = \chi(\Upsilon) = -\beta_w,$$

where the first equality is by Lemma 6.2, the second equality is by Lemma 6.4, and the fourth equality is by (6-4). □

7. Open questions

This paper raises quite a few questions and directions for future research, and we gather the main ones here. As above, $\mathcal{A} = K[\boldsymbol{F}]$ and $\pi_q(w)$ is the q-primitivity rank of $w \in \boldsymbol{F}$ (see Definition 1.5).

Expected number of fixed vectors. As stated in Conjecture 1.6, is it true that for every $w \in \boldsymbol{F}$, we have $\mathbb{E}_w[\mathrm{fix}] = 2 + \frac{|\mathrm{Crit}_q(w)|}{q^{N(\pi-1)}} + O\left(\frac{1}{q^{N\pi}}\right)$ where $\pi = \pi_q(w)$? If true, this would generalize Corollary 1.7 and yield that in free groups of arbitrary finite rank the words inducing the uniform measure on $\mathrm{GL}_N(K)$ for every N are precisely the primitive words — a result analogous to [Puder and Parzanchevski 2015, Theorem 1.1] dealing with S_N.

The q-primitivity rank. Recall Conjecture 1.9: is it true that $\pi_q(w) = \pi(w)$ for every $w \in \boldsymbol{F}$ and every prime power q? What is the value for a generic word (compare with [Puder 2015, Corollary 8.3] and [Kapovich 2022])? The Cohn–Lewin theorem applies to the free group algebra over an arbitrary field, not necessarily finite, and one can analogously define the K-primitivity rank of w for an arbitrary field K (and even for certain rings). Is it true that the K-primitivity rank is equal to $\pi(w)$ for every field K?

What about general elements of $\mathcal{A}$? One can define the primitivity rank $\pi_{\mathcal{A}}(f)$ of arbitrary $f \in \mathcal{A}$ as the rank of critical ideals, so $\pi_q(w) = \pi_{\mathcal{A}}(w-1)$ (and see the paragraph preceding Corollary 3.11). What are the possible values of $\pi_{\mathcal{A}}(f)$ for $f \in \mathcal{A}$? Does this number have any combinatorial meaning (à la Conjecture 1.6)?

The expected value of stable irreducible characters. Recall Conjecture 1.15 which says that for every stable irreducible character χ of $\mathrm{GL}_\bullet(K)$, $\mathbb{E}_w[\chi] = O((\dim \chi)^{1-\pi_q(w)})$. This conjecture should be quite difficult to tackle, as it is not even known in the somewhat simpler case of the symmetric group. It is more conceivable that one may be able to prove the weaker result that $\mathbb{E}_w[\chi] = O(q^{-N\cdot\pi_q(w)})$ for every nonpower w and every stable irreducible character of dimension $\Omega(q^{2N})$. This kind of result was proved for stable irreducible characters of $\{S_N\}_N$ [Hanany and Puder 2023, Corollary 1.7], for $\{U(N)\}_N$ [Brodsky 2024] and for $\{G \wr S_N\}_N$ for any finite group G [Shomroni 2023b]. (See also [Ernst-West et al. 2023, Appendix A] for further discussion and a more refined conjecture.)

Spectral gap in random Schreier graphs of $\mathbf{GL}_N(\boldsymbol{K})$**.** Part of the original motivation for studying word measures on $\mathrm{GL}_N(K)$ lies in questions regarding expansion and spectral gaps in random Schreier graphs of the groups $\mathrm{GL}_N(K)$ when K is fixed and $N \to \infty$. A recent milestone here is [Eberhard and Jezernik 2022]. Still, the following question is still open: Consider a random Schreier graph depicting the linear action of $\mathrm{GL}_N(K)$ on $K^N \setminus \{0\}$ with respect to two random generators. Do these graphs admit a uniform spectral gap with probability $\to 1$ as $N \to \infty$? If so, is the spectral gap optimal? It is plausible that the results and conjectures in this paper may contribute to obtaining such results, in a fashion similar to analogous proofs for Schreier graphs of S_N [Linial and Puder 2010; Puder 2015; Friedman and Puder 2023; Hanany and Puder 2023].

Limit distributions. Theorem 1.13 states that for w a nonpower, the distribution of the number of fixed vectors in a w-random element of $\mathrm{GL}_N(K)$ converges in distribution, as $N \to \infty$, to a limit distribution which is independent of w. Is this true for powers too? Is this true for an arbitrary stable class function in the ring $\mathcal{R}$ from page 2054? (This is known for S_N — see [Nica 1994, Theorem 1.1] and [Puder and Zimhoni 2024, Theorem 1.14] for a more general result about cycles of bounded length.)

Free group algebras. This paper gives rise to quite a few questions about the free group algebra $\mathcal{A}$. First, it is natural to guess that Corollary 5.2 can be generalized as follows: if $T \subseteq \mathrm{Cay}(\boldsymbol{F}, B)$ is a subtree and f, supported on T, is a primitive element of $I \trianglelefteq_T \mathcal{A}$, can $\{f\}$ be extended to a basis of I which is supported on T?

Recall Theorem 5.4 that when w is a nonpower, the only cyclic generators of the right $\mathcal{A}$-module $\mathcal{A}/(w-1)$ are images of unit elements of $\mathcal{A}$. Is this true for general subgroups of $\boldsymbol{F}$? Namely, let $H \le \boldsymbol{F}$ be a finitely generated subgroup which is not contained in any other subgroup of equal or smaller rank (in the language of [Puder 2014], this is $\pi(H) > \operatorname{rk} H$). Let $J_H \stackrel{\text{def}}{=} I_H\mathcal{A} = (\{h-1 \mid h \in H\})$ (see [Cohen 1972, Chapter 4]). Is it true that the only cyclic generators of the quotient $\mathcal{A}$-module $\mathcal{A}/J_H$ are images of unit elements of $\mathcal{A}$? This would be a Kaplansky-type result for such modules.

There are many other famous theorems and algorithms about free groups and their subgroups and we wonder if they have versions that apply to the free group algebra and its ideals. For example, is there an analog of Whitehead's cut vertex criterion which may detect efficiently whether a given element belongs to a free factor of a given ideal? See the recent survey [Delgado and Ventura 2022] giving a list of results about free groups and their subgroups using Stalling core graphs.

Appendix: The limit distribution of fix

Fix a nonpower $1 \neq w \in \boldsymbol{F}$. Recall that $\mathrm{fix}_{w,N}$ denotes the number of fixed vectors of a w-random element in $\mathrm{GL}_N(K)$. We explain why the method of moments is applicable for proving convergence in distribution for $\mathrm{fix}_{w,N}$, thus proving Theorem 1.13. We begin by recalling some basic definitions for the moment problem.

Given a sequence of real numbers $(m_n)_{n\ge 0}$ and an interval $I \subseteq \mathbb{R}$, a solution to the associated moment problem is a positive Borel measure θ supported on I with moments $\int_I x^n \, d\theta(x) = m_n$. When $I = \mathbb{R}$ (respectively, $I = [0, \infty)$), the problem is called a *Hamburger* (respectively, *Stieltjes*) moment problem. If a solution exists, the moment problem is said to be *solvable*. A solvable moment problem is further categorized by the number of solutions: if a unique solution exists, the moment problem is said to be *determinate* and otherwise it is called *indeterminate*, in which case there are infinitely many solutions since the set of solutions is convex.

The limiting measure of $\mathrm{fix}_{w,N}$ is a special case of a well-studied family of measures in the field of orthogonal polynomials. We next recall this family of measures, and then explain how previous analysis of the determinacy of its associated moment problems allows us to deduce the desired convergence in distribution.

Let $p \in (0, 1)$ and $a > 0$. The *Al-Salam Carlitz* polynomials of the second kind $V_n^{(a)}(x; p)$ (see [Chihara 1978, pp. 195–198; Koekoek et al. 2010, Section 14.24; Christiansen 2004, pp. 30-33]) are orthogonal with respect to the probability measure supported on the sequence $\{p^{-k}\}_{k\ge 0}$ with masses

$$w_{AC}(p^{-k}; a; p) = (ap; p)_\infty \frac{a^k p^{k^2}}{(p; p)_k (ap; p)_k}, \tag{A-1}$$

where $(x; y)_n = \prod_{j=0}^{n-1}(1 - xy^j)$ is the q-shifted factorial, or q-Pochhammer symbol, and $(x; y)_\infty = \prod_{j=0}^{\infty}(1 - xy^j)$.

Let q be a prime power. The limiting measure ν of $\mathrm{fix}_{w,N}$ is a special case of the family of measures (A-1) with parameters $p = q^{-1}$ and $a = 1$. Explicitly, $\nu = \sum_{k=0}^{\infty} w_{AC}(q^k; 1; q^{-1})\delta_{q^k}$. The n-th moment of ν is

equal to the number of linear subspaces of an n-dimensional vector space over a field with q elements (see [Fulman and Stanton 2016, Proposition 5.7] or [Chihara 1978, Equation 10.10]).

Let ν' be the pushforward of ν under the translation map $x \mapsto x-1$, i.e.,

$$\nu' = \sum_{k=0}^{\infty} w_{AC}(q^k; 1; q^{-1})\delta_{q^k-1}.$$

The measure ν' exhibits the interesting phenomenon of having its Hamburger moment problem be indeterminate while its Stieltjes moment problem is determinate (see [Berg and Valent 1994, Section 4]). Since the moments of a random variable Z determine the moments of $Z-1$ and vice versa, the pushforward map induced by $x \mapsto x-1$ forms a bijection between solutions to the moment problem associated to ν on $I = [1, \infty)$ and solutions to the Stieltjes moment problem associated to ν'. In particular, any measure supported on $[1, \infty)$ with the same moments as ν must be equal to ν.

Let ν_n be a sequence of Borel probability measures on $\mathbb{R}$ supported on $[1, \infty)$, and suppose that for every $k \in \mathbb{N}$ the k-th moment of ν_n converges as $k \to \infty$ to the k-th moment of ν, as Theorem 1.12 applied with $\mathcal{B} = I_k \in \mathrm{GL}_k(K)$ yields for $\mathrm{fix}_{w,N}$. We are now ready to deduce that ν_n converges weakly[14] to ν. The set of Borel probability measures on $\mathbb{R}$ equipped with the topology of weak convergence is metrizable (the Lévy metric, for example; see [Durrett 2019, Exercise 3.2.6]), and so it is enough to show that every subsequence of ν_n has a further subsequence converging weakly to ν. Let ν_{n_k} be such a subsequence. The convergence of the second moments implies that the sequence $(\nu_n)_{n\in\mathbb{N}}$ is tight [Durrett 2019, Theorem 3.2.14], and by Prokhorov's theorem [Billingsley 1999, Theorem 5.1], ν_{n_k} has a further subsequence $\nu_{n_{k_l}}$ converging weakly to some probability measure $\widetilde{\nu}$. The convergence of moments of ν_n to the moments of ν implies that $\widetilde{\nu}$ has the same sequence of moments as ν [Durrett 2019, Exercise 3.2.5]. Furthermore, using the Portmanteau theorem [Durrett 2019, Theorem 3.2.11] on the closed set $[1, \infty) \subseteq \mathbb{R}$, we get

$$\widetilde{\nu}([1, \infty)) \geq \limsup_{l\to\infty} \nu_{n_{k_l}}([1, \infty)) = \limsup_{l\to\infty} 1 = 1,$$

and so $\widetilde{\nu}$ must also be supported on $[1, \infty)$. The determinacy of the moment problem associated to ν on $I = [1, \infty)$ implies that $\widetilde{\nu} = \nu$, finishing the argument.

Acknowledgments

We thank Khalid Bou-Rabee for some prehistorical discussions on word measures on $\mathrm{GL}_2(K)$. We corresponded with Christian Berg on subjects around Theorem 1.13 and its proof, and we thank him for some very helpful pointers he gave us. We also thank George Bergman, Michael Magee and Yotam Shomroni for helpful discussions, and an anonymous referee for a thorough reading of the paper and many suggestions for small improvements. This research has received funding from the Israel Science Foundation: ISF grant 1071/16 and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 850956).

[14]The convergence in distribution of random variables is equivalent to the weak convergence of their measures.

References

[Berg and Valent 1994] C. Berg and G. Valent, "The Nevanlinna parametrization for some indeterminate Stieltjes moment problems associated with birth and death processes", *Methods Appl. Anal.* **1**:2 (1994), 169–209. MR Zbl

[Billingsley 1999] P. Billingsley, *Convergence of probability measures*, 2nd ed., Wiley, New York, 1999. MR Zbl

[Broder and Shamir 1987] A. Broder and E. Shamir, "On the second eigenvalue of random regular graphs", pp. 286–294 in 28*th Annual Symposium on Foundations of Computer Science* (Los Angeles, 1987), IEEE, Washington, DC, 1987. Zbl

[Brodskii 1984] S. D. Brodskii, "Equations over groups, and groups with one defining relation", *Sibirsk. Mat. Zh.* **25**:2 (1984), 84–103. In Russian; translated in *Siberian Math. J.* **25**:2 (1984), 235–251. MR Zbl

[Brodsky 2024] Y. Brodsky, "Word measures on unitary groups: improved bounds for small representations", *Int. Math. Res. Not.* (online publication May 2024), art. id. rnae100.

[Burns and Hale 1972] R. G. Burns and V. W. D. Hale, "A note on group rings of certain torsion-free groups", *Canad. Math. Bull.* **15**:3 (1972), 441–445. MR Zbl

[Chihara 1978] T. S. Chihara, *An introduction to orthogonal polynomials*, Math. Appl. **13**, Gordon & Breach, New York, 1978. MR Zbl

[Christiansen 2004] J. S. Christiansen, *Indeterminate moment problems within the Askey-scheme*, Ph.D. thesis, University of Copenhagen, 2004.

[Clay and Rolfsen 2016] A. Clay and D. Rolfsen, *Ordered groups and topology*, Grad. Stud. Math. **176**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[Cohen 1972] D. E. Cohen, *Groups of cohomological dimension one*, Lecture Notes in Math. **245**, Springer, 1972. MR Zbl

[Cohn 1964] P. M. Cohn, "Free ideal rings", *J. Algebra* **1**:1 (1964), 47–69. MR Zbl

[Collins and Śniady 2006] B. Collins and P. Śniady, "Integration with respect to the Haar measure on unitary, orthogonal and symplectic group", *Comm. Math. Phys.* **264**:3 (2006), 773–795. MR Zbl

[Delgado and Ventura 2022] J. Delgado and E. Ventura, "A list of applications of Stallings automata", *Trans. Comb.* **11**:3 (2022), 181–235. MR Zbl

[Durrett 2019] R. Durrett, *Probability: theory and examples*, 5th ed., Cambridge Ser. Statist. Probab. Math. **49**, Cambridge Univ. Press, 2019. MR Zbl

[Eberhard and Jezernik 2022] S. Eberhard and U. Jezernik, "Babai's conjecture for high-rank classical groups with random generators", *Invent. Math.* **227**:1 (2022), 149–210. MR Zbl

[Ernst-West 2019] D. West, *Word measures on* $\mathrm{GL}_n(\mathbb{F}_q)$, master's thesis, Tel Aviv University, 2019.

[Ernst-West et al. 2023] D. Ernst-West, D. Puder, and M. Seidel, "Stable q-primitivity rank and stable characters of $\mathrm{GL}_\bullet(q)$", 2023. Appendix to D. Puder and Y. Shomroni, "Stable invariants and their role in word measures on groups", arXiv:2311.17733.

[Ernst-West et al. 2024] D. Ernst-West, D. Puder, and Y. Shomroni, "The ring of stable characters over $\mathrm{GL}_\bullet(q)$", preprint, 2024. arXiv 2409.16571

[Friedman and Puder 2023] J. Friedman and D. Puder, "A note on the trace method for random regular graphs", *Israel J. Math.* **256**:1 (2023), 269–282. MR Zbl

[Fulman and Stanton 2016] J. Fulman and D. Stanton, "On the distribution of the number of fixed vectors for the finite classical groups", *Ann. Comb.* **20**:4 (2016), 755–773. MR Zbl

[Gan and Watterlond 2018] W. L. Gan and J. Watterlond, "A representation stability theorem for VI-modules", *Algebr. Represent. Theory* **21**:1 (2018), 47–60. MR Zbl

[Gardam 2021] G. Gardam, "A counterexample to the unit conjecture for group rings", *Ann. of Math.* (2) **194**:3 (2021), 967–979. MR Zbl

[Hanany and Puder 2023] L. Hanany and D. Puder, "Word measures on symmetric groups", *Int. Math. Res. Not.* **2023**:11 (2023), 9221–9297. MR Zbl

[Hanany et al. 2020] L. Hanany, C. Meiri, and D. Puder, "Some orbits of free words that are determined by measures on finite groups", *J. Algebra* **555** (2020), 305–324. MR Zbl

[Hog-Angeloni 1990] C. Hog-Angeloni, "A short topological proof of Cohn's theorem", pp. 90–95 in *Topology and combinatorial group theory*, edited by P. Latiolais, Lecture Notes in Math. **1440**, Springer, 1990. MR Zbl

[Howie 1982] J. Howie, "On locally indicable groups", *Math. Z.* **180**:4 (1982), 445–461. MR Zbl

[Kapovich 2022] I. Kapovich, "Primitivity rank for random elements in free groups", *J. Group Theory* **25**:5 (2022), 823–835. MR Zbl

[Karrass et al. 1960] A. Karrass, W. Magnus, and D. Solitar, "Elements of finite order in groups with a single defining relation", *Comm. Pure Appl. Math.* **13**:1 (1960), 57–66. MR Zbl

[Koekoek et al. 2010] R. Koekoek, P. A. Lesky, and R. F. Swarttouw, *Hypergeometric orthogonal polynomials and their q-analogues*, Springer, 2010. MR Zbl

[Lewin 1969] J. Lewin, "Free modules over free algebras and free group algebras: the Schreier technique", *Trans. Amer. Math. Soc.* **145** (1969), 455–465. MR Zbl

[Linial and Puder 2010] N. Linial and D. Puder, "Word maps and spectra of random graph lifts", *Random Structures Algorithms* **37**:1 (2010), 100–135. MR Zbl

[Louder and Wilton 2022] L. Louder and H. Wilton, "Negative immersions for one-relator groups", *Duke Math. J.* **171**:3 (2022), 547–594. MR Zbl

[Magee 2022] M. Magee, "Random unitary representations of surface groups, I: Asymptotic expansions", *Comm. Math. Phys.* **391**:1 (2022), 119–171. MR Zbl

[Magee and Puder 2019] M. Magee and D. Puder, "Matrix group integrals, surfaces, and mapping class groups, I: $\mathcal{U}(n)$", *Invent. Math.* **218**:2 (2019), 341–411. MR Zbl

[Magee and Puder 2021] M. Magee and D. Puder, "Surface words are determined by word measures on groups", *Israel J. Math.* **241**:2 (2021), 749–774. MR Zbl

[Magee and Puder 2023] M. Magee and D. Puder, "The asymptotic statistics of random covering surfaces", *Forum Math. Pi* **11** (2023), art. id. e15. MR Zbl

[Magee and Puder 2024] M. Magee and D. Puder, "Matrix group integrals, surfaces, and mapping class groups, II: $\mathrm{O}(n)$ and $\mathrm{Sp}(n)$", *Math. Ann.* **388**:2 (2024), 1437–1494. MR Zbl

[Mingo et al. 2007] J. A. Mingo, P. Śniady, and R. Speicher, "Second order freeness and fluctuations of random matrices, II: Unitary random matrices", *Adv. Math.* **209**:1 (2007), 212–240. MR Zbl

[Nica 1994] A. Nica, "On the number of cycles of given length of a free word in several random permutations", *Random Structures Algorithms* **5**:5 (1994), 703–730. MR Zbl

[Puder 2014] D. Puder, "Primitive words, free factors and measure preservation", *Israel J. Math.* **201**:1 (2014), 25–73. MR Zbl

[Puder 2015] D. Puder, "Expansion of random graphs: new proofs, new results", *Invent. Math.* **201**:3 (2015), 845–908. MR Zbl

[Puder and Parzanchevski 2015] D. Puder and O. Parzanchevski, "Measure preserving words are primitive", *J. Amer. Math. Soc.* **28**:1 (2015), 63–97. MR Zbl

[Puder and Zimhoni 2024] D. Puder and T. Zimhoni, "Local statistics of random permutations from free products", *Int. Math. Res. Not.* **2024**:5 (2024), 4242–4300. MR Zbl

[Putman and Sam 2017] A. Putman and S. V. Sam, "Representation stability and finite linear groups", *Duke Math. J.* **166**:13 (2017), 2521–2598. MR Zbl

[Rosenmann 1993] A. Rosenmann, "An algorithm for constructing Gröbner and free Schreier bases in free group algebras", *J. Symbolic Comput.* **16**:6 (1993), 523–549. MR Zbl

[Rosenmann and Rosset 1994] A. Rosenmann and S. Rosset, "Ideals of finite codimension in free algebras and the fc-localization", *Pacific J. Math.* **162**:2 (1994), 351–371. MR Zbl

[Rădulescu 2006] F. Rădulescu, "Combinatorial aspects of Connes's embedding conjecture and asymptotic distribution of traces of products of unitaries", pp. 197–205 in *Operator theory* 20 (Timişoara, Romania, 2004), edited by K. R. Davidson et al., Theta Ser. Adv. Math. **6**, Theta, Bucharest, 2006. MR Zbl

[Shomroni 2023a] Y. Shomroni, "Word measures on wreath products, I", preprint, 2023. arXiv 2305.11285

[Shomroni 2023b] Y. Shomroni, "Word measures on wreath products, II", preprint, 2023. arXiv 2311.11316

[Umirbaev 1994] U. U. Umirbaev, "Primitive elements of free groups", *Uspekhi Mat. Nauk* **49**:2(296) (1994), 175–176. In Russian; translated in *Russian Math. Surv.* **49**:2 (1994), 184–185. MR Zbl

[Weinbaum 1972] C. M. Weinbaum, "On relators and diagrams for groups with one defining relation", *Illinois J. Math.* **16**:2 (1972), 308–322. MR Zbl

Communicated by Michael J. Larsen
Received 2022-12-06 Revised 2023-04-21 Accepted 2023-11-27

daniellewest@mail.tau.ac.il *School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel*

doronpuder@gmail.com *School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel*

matanseidel@gmail.com *School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel*

ALGEBRA AND NUMBER THEORY 18:11 (2024)
https://doi.org/10.2140/ant.2024.18.2091

The distribution of large quadratic character sums and applications

Youness Lamzouri

Dedicated to Andrew Granville on the occasion of his 60th birthday

We investigate the distribution of the maximum of character sums over the family of primitive quadratic characters attached to fundamental discriminants $|d| \le x$. In particular, our work improves results of Montgomery and Vaughan, and gives strong evidence that the Omega result of Bateman and Chowla for quadratic character sums is optimal. We also obtain similar results for real characters with prime discriminants up to x, and deduce the interesting consequence that almost all primes with large Legendre symbol sums are congruent to 3 modulo 4. Our results are motivated by a recent work of Bober, Goldmakher, Granville and Koukoulopoulos, who proved similar results for the family of nonprincipal characters modulo a large prime. However, their method does not seem to generalize to other families of Dirichlet characters. Instead, we use a different and more streamlined approach, which relies mainly on the quadratic large sieve. As an application, we consider a question of Montgomery concerning the positivity of sums of Legendre symbols.

1. Introduction

Character sums play a central role in modern number theory through their numerous applications in the study of various arithmetic, analytic, algebraic and geometric objects. One important and basic example is that of quadratic Dirichlet characters, which include the Legendre symbol. The study of such characters and related sums has a long and rich history stretching back to the work of Gauss on binary quadratic forms. Let χ be a Dirichlet character modulo q. One quantity that was extensively studied over the past century is

$$M(\chi) := \max_{t \le q} \left| \sum_{n \le t} \chi(n) \right|.$$

In 1918, Pólya and Vinogradov independently proved that

$$M(\chi) \ll \sqrt{q} \log q.$$

MSC2020: primary 11L40, 11N64; secondary 11K65.
Keywords: character sums, Dirichlet L-functions, Pólya–Vinogradov inequality, quadratic large sieve, Kronecker symbol, Legendre symbol, positivity of partial sums.

On the other hand, an easy argument (based on applying Parseval's theorem to the Pólya Fourier series (2-1) attached to χ) shows that $M(\chi) \gg \sqrt{q}$ for all primitive characters modulo q. Though one can establish the Pólya–Vinogradov inequality using only basic Fourier analysis, improving on it has proved to be a very difficult problem, and resisted substantial progress outside of special cases. We should also note that any such improvement would have important consequences on several important quantities in analytic number theory, including class numbers of quadratic fields, short character sums, and the least quadratic nonresidue (see, for example, [Bober and Goldmakher 2016; Granville and Mangerel 2023; Mangerel 2020]). Granville and Soundararajan [2007] made an important breakthrough by showing that the Pólya–Vinogradov inequality can substantially be improved for characters of a fixed odd order. Further improvements to this case were obtained by Goldmakher [2012], and Lamzouri and Mangerel [2022].

Since $\sqrt{q} \ll M(\chi) \ll \sqrt{q}\log q$ for all primitive characters modulo q, a natural question is to determine the maximal order of $M(\chi)$. Assuming the generalized Riemann hypothesis (GRH), Montgomery and Vaughan [1977] proved that $M(\chi) \ll \sqrt{q}\log\log q$. This turns out to be optimal (up to a constant factor) in view of an older result of Paley [1932] who proved the existence of an infinite family of quadratic characters for which $M(\chi)$ is that large. Granville and Soundararajan [2007] refined Montgomery and Vaughan's conditional result and showed that

$$M(\chi) \le (2C_\chi + o(1))\sqrt{q}\log\log q, \tag{1-1}$$

for all primitive characters χ modulo q, where $C_\chi = e^\gamma/\pi$ if χ is odd, and $C_\chi = e^\gamma/(\sqrt{3}\pi)$ if χ is even. On the other hand, Paley's result was refined by Bateman and Chowla [1950], who proved the existence of an infinite sequence of moduli q, and primitive quadratic characters $\chi \bmod q$, such that

$$M(\chi) \ge \Big(\frac{e^\gamma}{\pi} + o(1)\Big)\sqrt{q}\log\log q. \tag{1-2}$$

This result was extended to the family of primitive characters modulo a large prime q by several authors (see, for example, Theorem 3 of [Granville and Soundararajan 2007]). Finally, we should note that Granville and Soundararajan [2007] conjectured that Bateman and Chowla's Omega result should correspond to the true extreme values of $M(\chi)$, namely that

$$M(\chi) \le (C_\chi + o(1))\sqrt{q}\log\log q, \tag{1-3}$$

for all primitive characters χ.

1A. ***The distribution of character sums.*** In view of (1-1), (1-2) and (1-3) it is natural to renormalize $M(\chi)$ by defining

$$m(\chi) := \frac{e^{-\gamma}\pi}{\sqrt{q}} M(\chi).$$

Montgomery and Vaughan [1979] were the first to study the distribution of $m(\chi)$ over families of Dirichlet characters. In particular, they showed that $m(\chi)$ is bounded (and hence $M(\chi) \ll \sqrt{q}$) for most characters.

Let q be a large prime and

$$\Phi_q(\tau) := \frac{1}{\varphi(q)}|\{\chi \neq \chi_0 \pmod q : m(\chi) > \tau\}|,$$

where $\varphi(q)$ is Euler's totient function. It follows from [Montgomery and Vaughan 1979] that

$$\Phi_q(\tau) \ll_A \tau^{-A},$$

for any constant $A \geq 1$. This estimate was improved by Bober and Goldmakher [2013] for fixed τ, and subsequently by Bober, Goldmakher, Granville and Koukoulopoulos [Bober et al. 2018] who showed that uniformly for $2 \leq \tau \leq \log\log q - M$ (where $M \geq 4$ is a parameter) we have

$$\exp\Big(-\frac{e^{\tau+A_0-\eta}}{\tau}\big(1+O(E_1(\tau,M))\big)\Big) \leq \Phi_q(\tau) \leq \exp\Big(-\frac{e^{\tau-2-\eta}}{\tau}\big(1+O(E_2(\tau))\big)\Big), \tag{1-4}$$

where $E_1(\tau, M) = (\log\tau)^2/\sqrt{\tau} + e^{-M/2}$, $E_2(\tau) = (\log\tau)/\tau$, $\eta := e^{-\gamma}\log 2$, and $A_0 = 0.088546\ldots$ is an explicit constant which can be expressed as a sum of integrals over the modified Bessel function of the first kind. In particular, this result gives strong evidence to the Granville–Soundararajan conjecture (1-3) for odd primitive characters modulo a large prime q.

Although the family of quadratic characters was the first for which large character sums were exhibited by Paley [1932], and then later by Bateman and Chowla [1950], no such distribution results are known in this case. In fact, the only known result for real characters is a result of Montgomery and Vaughan [1979] who showed that $\max_t\big|\sum_{n\leq t}\big(\frac{n}{p}\big)\big| \ll \sqrt{p}$ for most primes $p \leq x$. The main reason which explains why the results of [Bober and Goldmakher 2013; Bober et al. 2018] do not carry over to this setting is the fact that they rely heavily on the orthogonality relations for characters modulo q. Indeed, the key ingredient in the proof of (1-4) is to estimate the off-diagonal terms when bounding large moments of the tail of the sum in Pólya's Fourier expansion (2-1) below, which the authors of [Bober et al. 2018] successfully achieved using intricate estimates involving divisor functions.

In this paper, we overcome this problem for the family of quadratic characters by using a different and more streamlined approach which relies on the quadratic large sieve. Before stating our results we need some notation. For a fundamental discriminant d we let $\chi_d(\cdot) = \big(\frac{d}{\cdot}\big)$ be the Kronecker symbol modulo $|d|$. It is useful here to consider the cases of positive and negative discriminants separately, since as Theorems 1.1 and 1.2 below show, the distribution of large values of $m(\chi_d)$ behaves differently in each case. The difference between these cases lies in the fact that the character χ_d is even if d is positive, and is odd if d is negative. Thus, in view of Conjecture (1-3) we expect the extreme values of $m(\chi_d)$ for $d > 0$ to be smaller by a factor of $\sqrt{3}$ compared to the case $d < 0$. We denote by $\mathcal{F}(x)$ the set of fundamental discriminants d such that $|d| \leq x$, and let $\mathcal{F}^+(x)$ (respectively $\mathcal{F}^-(x)$) be the subset of $\mathcal{F}(x)$ consisting of positive (respectively negative) discriminants. Then we have the following standard estimates (see, for example, Lemma 4.1 of [Granville and Soundararajan 2006])

$$|\mathcal{F}^{\pm}(x)| = \frac{3}{\pi^2}x + O_\varepsilon(x^{1/2+\varepsilon}).$$

Our goal is to estimate the distribution functions

$$\Psi_x^{\pm}(\tau) := \frac{1}{|\mathcal{F}^{\pm}(x)|}|\{d \in \mathcal{F}^{\pm}(x) : m(\chi_d) > \tau\}|,$$

uniformly for τ in the range $2 \le \tau \le (1+o(1))\log\log x$ in the case of $\Psi_x^-(\tau)$, and the range $2 \le \tau \le (1/\sqrt{3}+o(1))\log\log x$ in the case of $\Psi_x^+(\tau)$. Conjecture (1-3) implies that these ranges are best possible up to the term $o(\log\log x)$. Here and throughout we shall denote by $\log_k$ the k-th iteration of the natural logarithm. We prove the following results.

Theorem 1.1. *Let $\eta = e^{-\gamma}\log 2$, and x be a large real number. Uniformly for τ in the range $2 \le \tau \le \log_2 x + \log_5 x - \log_4 x - C$ (where $C > 0$ is a suitably large constant) we have*

$$\exp\left(-\frac{e^{\tau-\eta-B_0}}{\tau}\left(1+O\left(\frac{(\log\tau)^2}{\sqrt{\tau}}\right)\right)\right) \le \Psi_x^-(\tau) \le \exp\left(-\frac{e^{\tau-\eta-\log 2-2}}{\tau}\left(1+O\left(\frac{\log\tau}{\tau}\right)\right)\right),$$

where

$$B_0 = \int_0^1 \frac{\tanh y}{y}\,dy + \int_1^{\infty} \frac{\tanh y - 1}{y}\,dy = 0.8187\ldots \tag{1-5}$$

Theorem 1.2. *Let B_0 be the constant in Theorem 1.1. There exists positive constants C_1 and C_2 such that uniformly for τ in the range $2 \le \tau \le (\log_2 x + \log_5 x - \log_4 x - C_1)/\sqrt{3}$ we have*

$$\exp\left(-\frac{e^{\sqrt{3}\tau-B_0}}{\sqrt{3}\tau}\left(1+O\left(\frac{1}{\tau}\right)\right)\right) \le \Psi_x^+(\tau) \ll \exp\left(-\frac{e^{\sqrt{3}\tau}}{\tau^{C_2}}\right).$$

Remark 1.3. Theorems 1.1 and 1.2 (the latter is the analogue of Theorem 1.3 of [Bober et al. 2018]) give strong evidence to the Granville–Soundararajan conjecture (1-3) for the family of quadratic characters. Moreover, a direct consequence of these results is the fact that almost all fundamental discriminants $|d| \le x$ for which $M(\chi_d)$ is large are negative.

Remark 1.4. The lower bounds in (1-4) and Theorems 1.1 and 1.2 are consequences of the works of Granville and Soundararajan [2003; 2006] on the distribution of $|L(1,\chi)|$ (over nonprincipal characters modulo q) in the case of (1-4), and $L(1,\chi_d)$ (over fundamental discriminants $|d| \le x$) in the case of Theorems 1.1 and 1.2 (see Section 2 for more details). Moreover, note that we obtain a slightly different constant in the upper bound of Theorem 1.1 compared to (1-4). We believe that this is caused by the different nature of the family of quadratic characters. Indeed in our case, the difference between the constants in the upper and lower bounds of Theorem 1.1 is $2+\log 2 - B_0 \approx 1.8744$, which is slightly smaller than the analogous difference $2 + A_0 \approx 2.0885$ for the family of nonprincipal characters modulo a large prime q in (1-4). Finally, we note that by using our approach, we can derive an easier proof of the upper bound of (1-4) in the range $2 \le \tau \le \log_2 q + \log_5 q - \log_4 q - C$. This is achieved by following the exact same argument of the proof of the upper bound of Theorem 1.1, and replacing the quadratic

large sieve inequalities of Heath-Brown and Elliott (see Lemma 4.1 below) by the following large sieve estimate of [Montgomery 1971, Theorem 6.2]

$$\sum_{\chi \bmod q} \left| \sum_{n \le N} a_n \chi(n) \right|^2 \ll (q+N) \sum_{n \le N} |a_n|^2,$$

which holds for an arbitrary complex sequence $\{a_n\}_{n \ge 1}$, and all integers $q \ge 2$. Although our approach gives a slightly smaller range of τ in this case, it has the advantage of extending the upper bound of (1-4) to all moduli q.

Remark 1.5. Using our work we can show that the structure results for large character sums obtained by Bober, Goldmakher, Granville and Koukoulopoulos in Section 2 of [Bober et al. 2018] for the family of nonprincipal characters modulo a large prime q, hold verbatim for the family of quadratic characters χ_d attached to fundamental discriminants $|d| \le x$. Since these results are technical, and the statements are exactly the same, we prefer to not state them here and refer the reader to the exact statements in [Bober et al. 2018]. The proofs follow along the same lines of [Bober et al. 2018] by using the auxiliary lemmas therein which hold for all primitive characters (and are derived using the "pretentious" theory of character sums developed by Granville and Soundararajan [2007]), and replacing the ingredients of the proof of Theorem 1.1 in [Bober et al. 2018] by those of the proof of Theorem 1.1 of our paper.

1B. ***Analogous results for prime discriminants.*** Using our approach we establish similar results to Theorems 1.1 and 1.2 over prime discriminants. The analogous lower bounds are direct consequences of newly established results on the distribution of $L\big(1, \big(\frac{\cdot}{p}\big)\big)$, which we shall describe in the next section. Furthermore, the analogous upper bounds will be obtained using the same methods of proofs of Theorems 1.1 and 1.2, together with the large sieve inequality of Montgomery and Vaughan [1979] for prime discriminants (see Lemma 5.1 below). Since the Legendre symbol modulo p is even if $p \equiv 1 \bmod 4$, and is odd if $p \equiv 3 \bmod 4$, we shall consider the cases of primes congruent to 1 and 3 modulo 4 separately. For $a \in \{1, 3\}$, let $\Psi_{x,a}^{\text{prime}}(\tau)$ be the proportion of primes $p \le x$ such that $p \equiv a \bmod 4$ and $m\big(\big(\frac{\cdot}{p}\big)\big) > \tau$.

Theorem 1.6. *Let η and B_0 be the constants in Theorem 1.1. There exists positive constants C_1, C_2 such that*:

1. *Uniformly in the range $2 \le \tau \le \log_2 x + \log_5 x - \log_4 x - C_1$ we have*

$$\Psi_{x,3}^{\text{prime}}(\tau) \le \exp\left(-\frac{e^{\tau-\eta-\log 2-2}}{\tau}\left(1+O\left(\frac{\log \tau}{\tau}\right)\right)\right).$$

2. *Uniformly in the range $2 \le \tau \le (\log_2 x)/2 - 2\log_3 x$ we have*

$$\Psi_{x,3}^{\text{prime}}(\tau) \ge \exp\left(-\frac{e^{\tau-\eta-B_0}}{\tau}(1+o(1))\right). \tag{1-6}$$

3. *Uniformly in the range $(\log_2 x)/2 - 2\log_3 x \le \tau \le \log_2 x - \log_3 x - C_1$ we have*

$$\Psi_{x,3}^{\text{prime}}(\tau) \ge \exp(-C_2 \tau e^{\tau}).$$

Theorem 1.7. *Let B_0 be the constant in Theorem 1.1. There exists positive constants C_1, C_2 and C_3 such that*:

1. *Uniformly in the range $2 \le \tau \le (\log_2 x - \log_3 x - C_1)/\sqrt{3}$ we have*

$$\exp(-C_3\tau e^{\sqrt{3}\tau}) \le \Psi_{x,1}^{\text{prime}}(\tau) \ll \exp\left(-\frac{e^{\sqrt{3}\tau}}{\tau^{C_2}}\right).$$

2. *Moreover, uniformly in the smaller range $2 \le \tau \le \big((\log_2 x)/2 - 2\log_3 x\big)/\sqrt{3}$ we have the improved lower bound*

$$\Psi_{x,1}^{\text{prime}}(\tau) \ge \exp\left(-\frac{e^{\sqrt{3}\tau - B_0}}{\sqrt{3}\tau}(1+o(1))\right). \tag{1-7}$$

Remark 1.8. An interesting consequence of Theorems 1.6 and 1.7 is that almost all primes $p \le x$ for which $M\big(\big(\frac{\cdot}{p}\big)\big)$ is large are congruent to 3 modulo 4. Moreover, we note that conditionally on the GRH, the lower bound (1-6) (respectively (1-7)) holds in the extended range $2 \le \tau \le \log_2 x - 2\log_3 x - C_4$ (respectively $2 \le \tau \le (\log_2 x - 2\log_3 x - C_4)/\sqrt{3}$), for some positive constant C_4. See Remark 6.1 below for a justification of this fact.

Remark 1.9. Theorem 1.4 of [Bober et al. 2018] establishes the existence of a limiting distribution for $m(\chi)$, as χ varies over nonprincipal characters modulo a large prime q, when $q \to \infty$. In [Hussain and Lamzouri 2023], we established an analogous result for $M\big(\big(\frac{\cdot}{p}\big)\big)$ as p varies over the primes in a large dyadic interval $[Q, 2Q]$ and $Q \to \infty$.

1C. ***The distribution of $L\big(1, \big(\frac{\cdot}{p}\big)\big)$.*** In order to prove the lower bounds of Theorems 1.6 and 1.7 we need estimates on the distribution of $L\big(1, \big(\frac{\cdot}{p}\big)\big)$, as p varies over the primes. It turns out that this case is harder than the case of the bigger family of primitive quadratic characters attached to fundamental discriminants d, which was investigated by Granville and Soundararajan [2003] (see the precise statement of their result in Theorem 2.2 below). By the law of quadratic reciprocity, this is due to the fact that current bounds on character sums over primes are much weaker than analogous bounds for character sums over the integers.[1] In particular, such bounds are heavily affected by the possible existence of a Landau–Siegel exceptional discriminant.

Joshi [1970] extended Littlewood's Omega result [1928], by establishing the existence of infinitely many primes p such that

$$L\left(1, \left(\frac{\cdot}{p}\right)\right) \ge (e^{\gamma} + o(1)) \log\log p.$$

[1] Bounds for character sums over primes are ultimately connected to the distribution of primes in arithmetic progressions. The weakness of such bounds explains for example why the strongest version of the prime number theorem for arithmetic progressions, namely the Siegel–Walfisz theorem, only holds for very small moduli.

We improve on this result by obtaining estimates for the proportion of primes $p \le x$ such that $L\big(1, \big(\frac{\cdot}{p}\big)\big) > e^{\gamma}\tau$ uniformly for τ in the range $2 \le \tau \le (1 - o(1)) \log\log p$, which is believed to be best possible up to the factor $o(\log\log p)$ (see for example the conjectures of Montgomery and Vaughan [1999] and Granville and Soundararajan [2003]). For $\tau > 0$ and $a \in \{1, 3\}$ we define

$$F_{x,a}(\tau) := \frac{2}{\pi(x)}\Big|\Big\{p \le x : p \equiv a \bmod 4,\ L\Big(1, \Big(\frac{\cdot}{p}\Big)\Big) > e^{\gamma}\tau\Big\}\Big|.$$

Theorem 1.10. *Let $a \in \{1, 3\}$ and x be large. In the range $2 \le \tau \le (\log_2 x)/2 - 2\log_3 x$ we have*

$$F_{x,a}(\tau) = \exp\Big(-\frac{e^{\tau - B_0}}{\tau}(1 + o(1))\Big). \tag{1-8}$$

Moreover, there exists positive constants $C_1, C_2 > 0$ such that in the range $(\log_2 x)/2 - 2\log_3 x \le \tau \le \log_2 x - \log_3 x - C_2$ we have

$$\exp(-C_1\tau e^{\tau}) \le F_{x,a}(\tau) \le \exp\Big(-\frac{e^{\tau + \log 2 - 2}}{\tau}\Big(1 + O\Big(\frac{\log\tau}{\tau}\Big)\Big)\Big). \tag{1-9}$$

Furthermore, the same estimates hold for the proportion of primes $p \le x$ such that $p \equiv a \bmod 4$ and $L\big(1, \big(\frac{\cdot}{p}\big)\chi_{-3}\big) > (2e^{\gamma}/3)\tau$, where χ_{-3} is the nonprincipal character modulo 3.

Remark 1.11. One can compare our results with those of [Granville and Soundararajan 2003] for the distribution of $L(1, \chi_d)$ over fundamental discriminants $|d| \le x$ (see Theorem 2.2 below). In this case, the same estimate (1-8) holds for the proportion of fundamental discriminants $|d| \le x$ such that $L(1, \chi_d) > e^{\gamma}\tau$, uniformly for τ in the larger range $2 \le \tau \le \log_2 x + (1 - o(1)) \log_4 x$. We should also note that conditionally on GRH, the estimate (1-8) holds in the extended range $2 \le \tau \le \log_2 x - 2\log_3 x - C_2$ (for some positive constant C_2) by a result of Holmin, Jones, Kurlberg, McLeman and Petersen [Holmin et al. 2019] (see Remark 6.1 below). Finally, we note that in the case $a = 3$, our proof yields a better constant in the upper bound of (1-9) than if we just apply Theorem 1.6 directly (using the inequality (2-8) below).

To prove the precise estimate (1-8) in the smaller range $2 \le \tau \le (1/2 - o(1)) \log_2 x$, we use our previous work [Lamzouri 2017], where we established asymptotic formulas for the complex moments of $L\big(1, \big(\frac{\cdot}{p}\big)\big)$ involving a secondary term which is coming from a possible Landau–Siegel exceptional discriminant. Although we could not rule out that this term might be as large as the main term, we were able to show that it does not affect the leading term in the tail of the distribution of $L\big(1, \big(\frac{\cdot}{p}\big)\big)$. However, we note that the error term $o(1)$ inside the estimate (1-8) is not effective, due to the use of Siegel's theorem. The upper bound of (1-9) will be a consequence of Theorem 2.7 below, which is the key ingredient in the proofs of the upper bounds of Theorems 1.6 and 1.7. Finally, to obtain the lower bound of (1-9), we combine Theorem 2.7 with a strong form of Linnik's theorem established by Bombieri [1987] using his zero density estimates for Dirichlet L-functions.

1D. ***An application to a question of Montgomery.*** As an application of our results we consider the problem of the positivity of sums of the Legendre symbol. Let $p \geq 3$ be a prime and

$$S_p(t) = \sum_{n \leq t} \Big(\frac{n}{p}\Big).$$

The question of determining when $S_p(t)$ is positive was considered by several mathematicians including Fekete, Chowla, Montgomery and others. Since $S_p(t)$ is periodic of period p, one can renormalize the variable t and define for $\alpha \geq 0$ the function

$$f_p(\alpha) := \sum_{0 \leq n \leq \alpha p} \Big(\frac{n}{p}\Big),$$

which is periodic of period 1. One can extend f_p to all $\alpha \in \mathbb{R}$ by periodicity. Montgomery [1976] studied the following natural question: How frequently is $f_p(\alpha)$ positive for a prime $p \equiv 3 \bmod 4$? More precisely, he investigated the quantity

$$\lambda(p) := \mu\big(\{\alpha \in [0,1) : f_p(\alpha) > 0\}\big),$$

for such primes[2], where $\mu(\mathcal{C})$ denotes the Lebesgue measure of a measurable set $\mathcal{C}$. He proved in [Montgomery 1976] that for all primes $p \equiv 3 \bmod 4$ we have $\lambda(p) > 1/50$. He also established the existence of infinitely many such primes such that $\lambda(p) < 1/3 + \varepsilon$ for any fixed $\varepsilon > 0$. In her Master's thesis, Mehkari [2005] slightly improved the constant $1/50$ in the first result of Montgomery. She also ran extensive numerical computations which suggest that the value $1/3$ for his second result is optimal. Furthermore, she proved conditionally on GRH that $\lambda(p) \leq 0.764$ for a positive proportion of the primes $p \equiv 3 \bmod 4$, and that $\lambda(p) \geq 0.285$ for a positive proportion of the primes $p \equiv 3 \bmod 4$. Montgomery [1976] also wrote "the ideas found in our proof can also be used to show that there are infinitely many primes $p \equiv 3 \bmod 4$ such that $\lambda(p) > 1 - \varepsilon$." Using a different method, based on the proof of Theorem 1.6, we improve on these results by showing that for any $\varepsilon > 0$, both inequalities $\lambda(p) > 1 - \varepsilon$ and $\lambda(p) < 1/3 + \varepsilon$ hold for a positive proportion of the primes $p \equiv 3 \bmod 4$. We also quantify these proportions in terms of ε and consider the question of uniformity by letting $\varepsilon \to 0$ slowly as a function of x, if we vary over the primes $p \leq x$ such that $p \equiv 3 \bmod 4$.

Theorem 1.12. *Let $\nu > 0$ be a small fixed constant. Let x be large and*

$$1 < T \leq \exp\big((1-\nu)\log_2 x \log_3 x/(\log_4 x)\big)$$

be a real number. The number of primes $p \leq x$ with $p \equiv 3 \bmod 4$ and such that $\lambda(p) > 1 - 1/T$ is

$$\gg \pi(x) \exp\bigg(-\exp\bigg(\frac{\log T \log_3 T}{\log_2 T}(1 + o(1))\bigg)\bigg). \tag{1-10}$$

[2]Note that f_p is even if $p \equiv 3 \bmod 4$ and is odd if $p \equiv 1 \bmod 4$. The analogous question in the latter case becomes: how often is $f_p(\alpha) > 0$ for $0 < \alpha < 1/2$? We shall only consider the former case in this paper since our methods can be extended to handle the case of primes $p \equiv 1 \bmod 4$.

In particular, this quantity is

$$\gg_\nu \pi(x)\exp(-T^\nu).$$

Theorem 1.13. *Let x be large and $T > 1$. There exists positive constants c_1, c_2, c_3 such that the number of primes $p \le x$ with $p \equiv 3 \bmod 4$ and such that $\lambda(p) < 1/3 + 1/T$ is at least*

$$\pi(x)\exp(-c_1 T^2(\log T)^3), \tag{1-11}$$

if $1 < T \le c_2(\log_2 x)^{1/2}/(\log_3 x)^2$, and is at least

$$\pi(x)\exp(-c_1 T^2(\log T)^5), \tag{1-12}$$

if $c_2(\log_2 x)^{1/2}/(\log_3 x)^2 \le T \le c_3(\log x)^{1/2}/(\log_2 x)^{5/2}$.

1E. ***Further applications.*** We should note that our method could be adapted to investigate the distribution of $M(\chi)$ over the family of primitive cubic Dirichlet characters with conductor up to x. We are planning to pursue this direction in a future paper. Moreover, we mention that the ideas of the proof of Theorem 1.6 are used in [Hussain and Lamzouri 2023], concerning the limiting distribution of character paths attached to the family of Legendre symbols modulo primes. Hussain [2022b] previously established a similar result for character paths attached to nonprincipal characters modulo a large prime q. Furthermore, in her PhD thesis [Hussain 2022a], she proved conditionally on GRH that character paths attached to the family of Legendre symbols converge in law (in the space of continuous functions) to a random Fourier series constructed using Rademacher random multiplicative functions. Our forthcoming work will establish this result unconditionally.

2. Outline and key ingredients of the proofs of Theorems 1.1, 1.2 and 1.6

2A. ***The overall strategy of the proof of Theorem 1.1.*** We first start by describing the strategy and key ideas of the proof of Theorem 1.1, since the method for proving Theorems 1.2, 1.6 and 1.7 is similar. In particular, it will be useful to compare the argument with [Bober et al. 2018]. The character sum $\sum_{n\le t}\chi_d(n)$ has a simple Fourier expansion first obtained by Pólya in the following quantitative form [Montgomery and Vaughan 2007, equation (9.19), p. 311]

$$\sum_{n\le t}\chi_d(n) = \frac{\mathcal{G}(\chi_d)}{2\pi i}\sum_{1\le|n|\le Z}\frac{\chi_d(n)(1-e(-nt/|d|))}{n} + O\left(1+\frac{|d|\log|d|}{Z}\right), \tag{2-1}$$

where $\mathcal{G}(\chi_d)$ is the Gauss sum attached to χ_d. Let $\delta := 1/100$, and define $\mathcal{F}^*(x)$ to be the set of fundamental discriminants $x^{1-\delta} \le |d| \le x$. We note that the proportion of those discriminants d with $|d| \le x^{1-\delta}$ is $\ll x^{-\delta}$, which is much smaller than the distribution function $\Psi_x^-(\tau)$ in the range $\tau \le \log\log x$. Hence, we will only focus on the fundamental discriminants $d \in \mathcal{F}^*(x)$. For these discriminants we have

$$m(\chi_d) = \frac{e^{-\gamma}}{2}\max_{\alpha\in[0,1)}\left|\sum_{1\le|n|\le Z}\frac{\chi_d(n)(1-e(\alpha n))}{n}\right| + O(x^{-\delta}), \tag{2-2}$$

where $Z = x^{21/40}$. We first describe a heuristic argument that will help us isolate the key ingredient in the proof of Theorem 1.1. This will also explain why we should expect the tail of the distribution of $m(\chi_d)$ to behave like $\exp(-e^\tau/\tau)$. A standard idea, which goes back to the work of Montgomery and Vaughan [1977] on exponential sums with multiplicative coefficients, is to split the sum on the right-hand side of (2-2) into two parts

$$\sum_{1\le|n|\le Z}\frac{\chi_d(n)(1-e(\alpha n))}{n}=\sum_{\substack{1\le|n|\le Z\\ P^+(n)\le y}}\frac{\chi_d(n)(1-e(\alpha n))}{n}+\sum_{\substack{1\le|n|\le Z\\ P^+(n)> y}}\frac{\chi_d(n)(1-e(\alpha n))}{n}, \tag{2-3}$$

and show that uniformly over α, the bulk of the distribution comes from the first part, over y-friable integers, with a suitable choice of the parameter y. Here $P^+(n)$ is the largest prime factor of n, and an integer n is called y-friable (or y-smooth) if $P^+(n)\le y$. To understand what choice of the parameter y we should make we observe that

$$\max_{\alpha\in[0,1)}\left|\sum_{\substack{1\le|n|\le Z\\ P^+(n)\le y}}\frac{\chi_d(n)(1-e(\alpha n))}{n}\right|\ll\sum_{\substack{n\ge1\\ P^+(n)\le y}}\frac{1}{n}\ll\log y \tag{2-4}$$

by Mertens' theorem. Hence, if the main part of the contribution to $m(\chi_d)>\tau$ is coming from y-friable integers, we should aim for a choice of y such that[3] $y\approx e^\tau$. Heuristically, for small y ($y\le\log x$ say) we can prescribe the values of $\chi_d(p)$ for $p\le y$ with probability[4] $2^{-\pi(y)}=\exp(-(\log 2+o(1))y/\log y)$ by the prime number theorem. Therefore, if $y\asymp e^\tau$ this probability looks like $\exp(-c_1e^\tau/\tau)$ for some positive constant c_1, which agrees with the statement of Theorem 1.1. Thus, in order for this heuristic to work, we need to efficiently control the second part in (2-3) uniformly over $\alpha\in[0,1)$. We achieve this in the following theorem, which is the key ingredient in the proofs of Theorems 1.1 and 1.2.

Theorem 2.1. *Let $h(n)$ be a completely multiplicative function such that $|h(n)|\le1$ for all n. Let x be large and put $Z=x^{21/40}$. There exists a constant $c>0$ such that for all real numbers $2\le y\le c\log x\log_4 x/(\log_3 x)$ and $1/\log y\le A\le4$, the number of fundamental discriminants $|d|\le x$ such that*

$$\max_{\alpha\in[0,1)}\left|\sum_{\substack{1\le n\le Z\\ P^+(n)> y}}\frac{\chi_d(n)h(n)e(n\alpha)}{n}\right|>e^\gamma A$$

is

$$\ll x\exp\left(-\frac{A^2y}{2\log y}\left(1+O\left(\frac{\log_2 y}{\log y}+\frac{\log_4 x}{A\log_2 x\log_3 x}\right)\right)\right).$$

[3]This is not completely correct, since the above argument shows that we rather have $y\approx e^{c\tau}$ for some constant $c>0$. However, it turns out that the optimal choice of c is $c=1$.

[4]This is correct in the range $y\le\log\log x$, as proved in Lemma 7.1 below. It also follows from GRH in the larger range $y\le c\log x\log\log x$ for some small constant $c>0$; see Theorem 13.5 of [Montgomery 1971].

Bober, Goldmakher, Granville and Koukoulopoulos [Bober et al. 2018] proved a similar result in the case of nonprincipal characters modulo a large prime q. However our proof differs from theirs as we shall now explain. Let $z = q^{21/40}$. In order to prove the analogue of Theorem 2.1 for the family of nonprincipal characters modulo q, the authors of [Bober et al. 2018] established nontrivial upper bounds for the $2k$-th moment of

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le z\\ P^+(n)>y}} \frac{\chi(n)e(n\alpha)}{n} \right| \tag{2-5}$$

for k roughly up to $\log q$. By decoupling the α from the character χ and expanding the $2k$-th moment of the inner sum in (2-5), one needs to control terms which have size as large as $z^k > q^{k/2}$ with k up to $\log q$. This was possible in the case of the family of nonprincipal characters modulo q thanks to the orthogonality relations of characters, which imply that the off-diagonal terms in these moments are given by $m \equiv n \mod q$ with $m \neq n$ and $m, n \le z^k$. The authors of [Bober et al. 2018] proceed to bound these off-diagonal terms using intricate estimates involving the k-th divisor function.

This argument is no longer valid for families of quadratic characters, since in this case only "quasi-orthogonality relations" are known (see, for example, Lemma 4.1 of [Granville and Soundararajan 2003]), which allow one to control terms up to size x^c for some $c < 1$, if we run over fundamental discriminants $|d| \le x$. To overcome this problem, we made the key observation that when $N < n < 2N$ and N is very large, one only needs to compute a small moment over the corresponding sum over the interval $[N, 2N]$, in order to show that this sum is small for most characters. More specifically, our approach consists of first splitting the sum in (2-5) (in the case of quadratic characters associated to fundamental discriminants $|d| \le x$) into two parts: the first over $n \le Y$ and the second over $Y < n \le Z$, where $Y = (\log x)^{W(x)}$ is relatively "small" (our method allows one to choose $W(x) = \log_3 x/(\log_4 x)$). We first bound the $2k$-th moments of the first sum over $n \le Y$ and show that only the diagonal terms contribute[5] to these moments if $k \le (\log x)/(3 \log Y)$ say. We then proceed to show that the second part over the large n's is very small for most characters. To this end, we split it in dyadic intervals $N \le n \le 2N$ and then bound the $2\ell_N$ moment of

$$\max_{\alpha\in[0,1)} \left| \sum_{N<n<2N} \frac{\chi(n)e(n\alpha)}{n} \right|$$

where ℓ_N is chosen such that N^{ℓ_N} is roughly of size x. This allows us to show that each of these parts is small for most characters using *tailored* moments according to the size of N. In the range $Y < N < x^\varepsilon$ we use (4-1) below, which is an easy version of the quadratic large sieve first due to Elliott, while in the remaining range $x^\varepsilon < N < Z$ we use the quadratic large sieve of Heath-Brown (see (4-2) below).

[5]We prefer to do this using an easy version of the quadratic large sieve (see (4-1) below), rather than employing the quasiorthogonality relations, in order to get a clean argument.

Deducing the upper bound of Theorem 1.1 from Theorem 2.1. Let $Z = x^{21/40}$ and $\delta = 1/100$. By (2-2) we have

$$m(\chi_d) \le \frac{e^{-\gamma}}{2} \max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le|n|\le Z\\ P^+(n)\le y}} \frac{\chi_d(n)(1-e(\alpha n))}{n} \right| + 2e^{-\gamma} \max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z\\ P^+(n)> y}} \frac{\chi_d(n)e(n\alpha)}{n} \right| + O(x^{-\delta}), \quad (2\text{-}6)$$

for any fundamental discriminant $-x \le d \le -x^{1-\delta}$. Since we do not have control over the first part over y-friable integers, we shall bound it using Corollary 3.5 of [Bober et al. 2018] which gives

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le|n|\le Z\\ P^+(n)\le y}} \frac{\chi_d(n)(1-e(\alpha n))}{n} \right| \le 2e^{\gamma}\log y + 2\log 2 + O\left(\frac{\log\log y}{\log y}\right). \quad (2\text{-}7)$$

We now set $y = e^{\tau - e^{-\gamma}\log 2 - 2B}$, where $B > 0$ is a parameter to be chosen. Combining this last estimate with (2-6) and (2-7) we deduce that

$$m(\chi_d) \le \tau - 2B + 2e^{-\gamma} \max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z\\ P^+(n)> y}} \frac{\chi_d(n)e(\alpha n)}{n} \right| + O\left(\frac{\log\tau}{\tau}\right).$$

Thus, the proportion of fundamental discriminants $-x \le d \le -x^{1-\delta}$ such that $m(\chi_d) > \tau$ is bounded by the proportion of fundamental discriminants $|d| \le x$ such that

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z\\ P^+(n)> y}} \frac{\chi_d(n)e(\alpha n)}{n} \right| > e^{\gamma}\left(B - C_0\frac{\log\tau}{\tau}\right),$$

for some suitably large constant $C_0 > 0$. Choosing $B = 1$ and appealing to Theorem 2.1 completes the proof. □

We now turn our attention to the lower bound of Theorem 1.1 which is easier. For $d < 0$ one has the identity (see Theorem 9.21 of [Montgomery and Vaughan 2007])

$$\sum_{n\le|d|/2} \chi_d(n) = (2-\chi(2))\frac{\mathcal{G}(\chi_d)}{i\pi} L(1,\chi_d),$$

which implies

$$m(\chi_d) \ge e^{-\gamma} L(1,\chi_d). \quad (2\text{-}8)$$

Therefore, one can immediately deduce a corresponding lower bound for the distribution function $\Psi_x^-(\tau)$ from the following result, which follows from a straightforward adaptation of the work of Granville and Soundararajan [2003] on the distribution of $L(1,\chi_d)$.

Theorem 2.2 (Granville–Soundararajan). *Let B_0 be the constant in Theorem 1.1, ψ be a character modulo some $b \in \{1, 3\}$, and $1 \le \tau \le \log_2 x + \log_4 x - 20 - M$ for some $M \ge 0$. Then we have*

$$\frac{1}{|\mathcal{F}^{\pm}(x)|}\left|\left\{d \in \mathcal{F}^{\pm}(x) : L(1, \chi_d\psi) > \frac{\phi(b)}{b}e^{\gamma}\tau\right\}\right| = \exp\left(-\frac{e^{\tau - B_0}}{\tau}\left(1 + O\left(\frac{1}{\tau} + e^{-e^M}\right)\right)\right).$$

One can do a little better by combining our Theorem 2.1 with the following result, which we extract from the proof of Theorem 1.2 of [Bober et al. 2018].

Theorem 2.3 (Bober, Goldmakher, Granville and Koukoulopoulos). *Let C be a positive constant. Let q be a large positive integer, and $y \le (\log q)^2$ be a large real number. Let χ be a primitive odd character modulo q such that*

$$|L(1, \chi)| > e^{\gamma}\log y - C \quad \text{and} \quad \max_{\alpha\in[0,1)}\left|\sum_{\substack{n\ge 1\\ P^+(n)>y}}\frac{\chi(n)e(n\alpha)}{n}\right| \le 1.$$

Then we have

$$m(\chi) > e^{-\gamma}\big(|L(1, \chi)| + \log 2\big) + O\left(\frac{(\log\log y)^2}{\sqrt{\log y}}\right).$$

Remark 2.4. We note that q is assumed to be prime in the statement of Theorem 1.2 of [Bober et al. 2018]. However, this is only used to show that many nonprincipal characters $\chi \bmod q$ satisfying the assumptions of Theorem 2.3 exist. Indeed, the proof of Theorem 2.3 (see the end of Section 4 of [Bober et al. 2018]) only uses estimates on exponential sums over y-friable integers [Bober et al. 2018, Lemmas 3.2 and 3.4] together with ideas of [Bober 2014] on averages of character sums to arbitrary moduli.

Deducing the lower bound of Theorem 1.1 from Theorems 2.1, 2.2 and 2.3. Let $Z = x^{21/40}$ and $\delta = 1/100$. By partial summation and the Pólya–Vinogradov inequality, it follows that

$$\max_{\alpha\in[0,1)}\left|\sum_{n>Z}\frac{\chi_d(n)e(n\alpha)}{n}\right| \ll \frac{\sqrt{|d|}\log|d|}{Z} \ll x^{-\delta}, \tag{2-9}$$

for any fundamental discriminant $|d| \le x$. Furthermore, by Lemma 3.2 of [Bober et al. 2018] we have

$$\max_{\alpha\in[0,1)}\left|\sum_{\substack{n>Z\\ P^+(n)\le y}}\frac{\chi_d(n)e(n\alpha)}{n}\right| \le \sum_{\substack{n>Z\\ P^+(n)\le y}}\frac{1}{n} \ll e^{-\sqrt{\log y}}, \tag{2-10}$$

for any real number $2 \le y \le (\log x)^2$. Combining these estimates implies that

$$\max_{\alpha\in[0,1)}\left|\sum_{\substack{n>Z\\ P^+(n)>y}}\frac{\chi_d(n)e(n\alpha)}{n}\right| \ll e^{-\sqrt{\log y}}, \tag{2-11}$$

for any fundamental discriminant $|d| \le x$ and any real number $2 \le y \le (\log x)^2$.

Let $C_1, C_2 > 0$ be suitably large constants and put $y = e^{\tau + C_1}$. If C_1 is large enough, then combining Theorems 2.1 and 2.2 we deduce that the number of fundamental discriminants $-x < d < 0$ such that

$$L(1, \chi_d) > e^{\gamma}\tau - \log 2 + C_1 \frac{(\log \tau)^2}{\sqrt{\tau}} \quad \text{and} \quad \max_{\alpha \in [0,1)} \Big| \sum_{\substack{1 \le n \le Z \\ P^+(n) > y}} \frac{\chi_d(n) e(n\alpha)}{n} \Big| \le 1 - \frac{C_2}{\tau}$$

is

$$\ge \exp\Big(-\frac{e^{\tau - \eta - B_0}}{\tau} \Big(1 + O\Big(\frac{(\log \tau)^2}{\sqrt{\tau}} \Big) \Big) \Big).$$

By (2-11) we deduce that for any such discriminant d we have

$$\max_{\alpha \in [0,1)} \Big| \sum_{\substack{n \ge 1 \\ P^+(n) > y}} \frac{\chi_d(n) e(n\alpha)}{n} \Big| \le 1,$$

and hence it follows from Theorem 2.3 that $m(\chi_d) > \tau$ if C_1 is suitably large. This completes the proof. □

2B. ***The case of positive discriminants: Proof of Theorem 1.2.*** Recall that χ_d is an even character if d is a positive fundamental discriminant. In this case, we shall appeal to the following structure result for even characters with large sums, which we extract from the proof of Theorem 2.3 of [Bober et al. 2018].

Theorem 2.5 (Bober, Goldmakher, Granville and Koukoulopoulos). *Let C be a positive constant. Let q be a large positive integer and $y \le (\log q)^2$ be a large real number. Let $\psi \bmod q$ be a primitive even character such that*

$$\max_{\alpha \in [0,1)} \Big| \sum_{\substack{n \ge 1 \\ P^+(n) \le y}} \frac{\psi(n) e(n\alpha)}{n} \Big| > \frac{e^{\gamma}}{\sqrt{3}} \log y - C \log\log y, \quad \text{and} \quad \max_{\alpha \in [0,1)} \Big| \sum_{\substack{n \ge 1 \\ P^+(n) > y}} \frac{\psi(n) e(n\alpha)}{n} \Big| \le 1.$$

Then we have

$$m(\psi) = \frac{e^{-\gamma}\sqrt{3}}{2} \big| L(1, \psi\chi_{-3}) \big| + O(\log\log y).$$

Remark 2.6. Here again we note that q is assumed to be prime in Theorem 2.3 of [Bober et al. 2018], but this is only used to show that many nonprincipal characters $\psi \bmod q$ satisfying the assumptions of Theorem 2.5 exist. Indeed, all of the ingredients of the proof of Theorem 2.5 hold for an arbitrary primitive even character ψ (see Section 8 of [Bober et al. 2018]), and are derived using the "pretentious" theory of character sums developed by Granville and Soundararajan [2007], along with estimates of Montgomery and Vaughan [1977] on exponential sums with multiplicative coefficients.

Deducing Theorem 1.2 from Theorems 2.1, 2.2 and 2.5. The lower bound follows immediately from Theorem 2.2 together with the standard inequality (see for example the beginning of Section 4 of [Bober et al. 2018])

$$M(\psi) \ge \frac{\sqrt{3q}}{2\pi} |L(1, \psi\chi_{-3})|, \tag{2-12}$$

which is valid for any primitive even character ψ modulo q.

We now prove the upper bound. First, it follows from (2-1) that for all fundamental discriminants $0 < d < x$ we have

$$m(\chi_d) = e^{-\gamma} \max_{\alpha\in[0,1)} \left| \sum_{1\le n\le Z} \frac{\chi_d(n)\sin(2\pi n\alpha)}{n} \right| + O(1), \tag{2-13}$$

since χ_d is even. Let $C_1, C_2 > 0$ be suitably large constants, $y = e^{\sqrt{3}\tau + C_1}$, and define $\mathcal{D}^+(x, y)$ to be the set of fundamental discriminants $0 < d \le x$ such that

$$m(\chi_d) > \tau \quad \text{and} \quad \max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z \\ P^+(n)>y}} \frac{\chi_d(n)e(n\alpha)}{n} \right| \le 1 - \frac{C_2}{\tau}.$$

By combining the lower bound of Theorem 1.2 with Theorem 2.1 we deduce that if C_1 is suitably large then

$$\frac{|\mathcal{D}^+(x, y)|}{|\mathcal{F}^+(x)|} = \Psi_x^+(\tau)(1 + o(1)). \tag{2-14}$$

Now, let $d \in \mathcal{D}^+(x, y)$. By (2-10) and (2-13) together with our assumption on d we have

$$m(\chi_d) = e^{-\gamma} \max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z \\ P^+(n)\le y}} \frac{\chi_d(n)\sin(2\pi n\alpha)}{n} \right| + O(1) = e^{-\gamma} \max_{\alpha\in[0,1)} \left| \sum_{\substack{n\ge 1 \\ P^+(n)\le y}} \frac{\chi_d(n)\sin(2\pi n\alpha)}{n} \right| + O(1).$$

Since $m(\chi_d) > \tau$ we deduce that

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{n\ge 1 \\ P^+(n)\le y}} \frac{\chi_d(n)e(\alpha n)}{n} \right| \ge e^{\gamma} m(\chi_d) + O(1) > \frac{e^{\gamma}}{\sqrt{3}} \log y - C_3,$$

for some positive constant C_3. Moreover, by (2-11) and our assumption on d we get

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{n\ge 1 \\ P^+(n)>y}} \frac{\chi_d(n)e(n\alpha)}{n} \right| \le 1 - \frac{C_2}{\tau} + O(e^{-\sqrt{\tau}}) \le 1,$$

if C_2 is suitably large. Thus χ_d satisfies the conditions of Theorem 2.5, which implies that

$$m(\chi_d) = \frac{e^{-\gamma}\sqrt{3}}{2} |L(1, \chi_d\chi_{-3})| + O(\log\tau).$$

Therefore, for all $d \in \mathcal{D}^+(x, y)$ we have

$$|L(1, \chi_d\chi_{-3})| > \frac{2e^{\gamma}}{\sqrt{3}}\tau - C_4 \log\tau,$$

for some constant $C_4 > 0$. Hence, it follows from Theorem 2.2 that

$$|\mathcal{D}^+(x, y)| \ll |\mathcal{F}^+(x)| \exp\left(-\frac{e^{\sqrt{3}\tau}}{\tau^{C_5}}\right),$$

for some constant $C_5 > 0$. Combining this estimate with (2-14) completes the proof. □

2C. ***The family of Legendre symbols: Outline of the proofs of Theorems 1.6 and 1.7.*** In order to prove the upper and lower bounds of Theorems 1.6 and 1.7, we shall follow the same strategy of the proofs of Theorems 1.1 and 1.2. To this end we establish the following key result, which is the analogue of Theorem 2.1 for prime discriminants. To shorten our notation, we let ψ_p denote the Legendre symbol modulo p throughout the remaining part of the paper.

Theorem 2.7. *Let $h(n)$ be a completely multiplicative function such that $|h(n)| \leq 1$ for all n. Let x be large and put $Z = x^{21/40}$. There exists a constant $c > 0$ such that for all real numbers $2 \leq y \leq c \log x \log_4 x/(\log_3 x)$ and $1/\log y \leq A \leq 4$, the number of primes $p \leq x$ such that*

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\leq n\leq Z\\ P^+(n)>y}} \frac{\psi_p(n)h(n)e(n\alpha)}{n} \right| > e^{\gamma} A$$

is

$$\ll \pi(x) \exp\left(-\frac{A^2 y}{2\log y}\left(1 + O\left(\frac{\log_2 y}{\log y} + \frac{\log_4 x}{A\log_2 x \log_3 x}\right)\right)\right).$$

We end this section by deducing Theorems 1.6 and 1.7 from this result along with Theorem 1.10 on the distribution of large values of $L(1, \psi_p)$.

Proof of Theorem 1.6 assuming Theorems 1.10 and 2.7. First, Part 1 of Theorem 1.6 can be derived along the same lines of the proof of the upper bound of Theorem 1.1 by replacing Theorem 2.1 by Theorem 2.7.

Now, the proof of Part 2 follows along the same lines of the proof of the lower bound of Theorem 1.1 by replacing Theorems 2.1 and 2.2 by Theorem 2.7 and the estimate (1-8) respectively. Finally, the proof of Part 3 follows from (1-9) together with the lower bound (2-8). □

Proof of Theorem 1.7 assuming Theorems 1.10 and 2.7. Part 2 and the lower bound of Part 1 of Theorem 1.7 follow immediately from Theorem 1.10 together with the lower bound (2-12). Finally, the upper bound of Part 1 can be derived along the same lines of the proof of the upper bound of Theorem 1.2 with the choice $y = \tau^2 e^{\sqrt{3}\tau + C_1}$ for some suitably large constant C_1, by using Theorem 2.5 and replacing Theorems 2.1 and 2.2 by Theorems 2.7 and 1.10 respectively. □

2D. ***The plan of the paper.*** The plan of the paper is as follows. In Section 3 we gather several preliminary results on sums of divisor functions and moments of Random multiplicative functions, which will shall use throughout the paper. Section 4 will be devoted to the proof of Theorem 2.1. Theorem 2.7 will be established in Section 5. In Section 6 we investigate the distribution of large values of $L(1, \psi_p)$ and prove Theorem 1.10. Finally, in Section 7, we investigate the positivity of sums of Legendre symbols and prove Theorems 1.12 and 1.13.

3. Sums of divisor functions and random multiplicative functions

In this section we gather together several preliminary results concerning sums of divisor functions, which are related to certain moments over random multiplicative functions. We let $\{\mathbb{X}(n)\}_{n\geq 1}$ be Rademacher

random multiplicative functions, that is $\{\mathbb{X}(p)\}_{p \text{ prime}}$ are I.I.D. random variables taking the values ± 1 with equal probability $1/2$, and we extend $\mathbb{X}(n)$ multiplicatively to all positive integers by setting $\mathbb{X}(1) = 1$ and $\mathbb{X}(n) = \prod_{p^\ell \| n} \mathbb{X}(p)^\ell$. We start with the following lemma.

Lemma 3.1. *Let k be a large real number. Then for any $0 \le \alpha \le \log_3 k/(2\log k)$ we have*

$$\sum_{n=1}^{\infty} \frac{d_k(n)^2}{n^{2-\alpha}} = \exp\big(O(k \log\log k)\big), \tag{3-1}$$

and

$$\sum_{n=1}^{\infty} \frac{d_k(n^2)}{n^{2-\alpha}} = \exp\big(O(k \log\log k)\big). \tag{3-2}$$

Moreover, for all $y > k$ we have

$$\sum_{P^-(n)>y} \frac{d_k(n^2)}{n^2} = \exp\left(O\left(\frac{k^2}{y\log y}\right)\right), \tag{3-3}$$

where here and throughout $P^-(n)$ denotes the smallest prime factor of n.

Proof. The bound (3-1) follows from Lemma 3.3 of [Lamzouri 2011]. Furthermore, the estimate (3-2) follows from (3-1) upon noting that $d_k(n^2) \le d_k(n)^2$. We now establish (3-3). Let $p > y$ be a prime number. Then we have

$$\begin{aligned}\mathbb{E}\left(\left(1-\frac{\mathbb{X}(p)}{p}\right)^{-k}\right) &= \frac{1}{2}\left(\left(1-\frac{1}{p}\right)^{-k} + \left(1+\frac{1}{p}\right)^{-k}\right)\\ &= \frac{1}{2}\left(\exp\left(\frac{k}{p} + O\left(\frac{k}{p^2}\right)\right) + \exp\left(-\frac{k}{p} + O\left(\frac{k}{p^2}\right)\right)\right)\\ &= 1 + O\left(\frac{k^2}{p^2}\right).\end{aligned} \tag{3-4}$$

Hence we derive

$$\begin{aligned}\sum_{P^-(n)>y} \frac{d_k(n^2)}{n^2} &= \mathbb{E}\left(\sum_{P^-(n)>y} \frac{d_k(n)\mathbb{X}(n)}{n}\right) = \prod_{p>y} \mathbb{E}\left(\left(1-\frac{\mathbb{X}(p)}{p}\right)^{-k}\right)\\ &= \exp\left(O\left(k^2 \sum_{p>y} \frac{1}{p^2}\right)\right) = \exp\left(O\left(\frac{k^2}{y\log y}\right)\right),\end{aligned}$$

as desired. □

In order to prove Theorems 2.1 and 2.7 we need a uniform bound for the moments

$$\mathcal{M}_y(k) := \mathbb{E}\left(\left|\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n}\right|^{2k}\right). \tag{3-5}$$

To this end we establish the following result.

Proposition 3.2. *Let $k \geq 2$ be a positive integer and $(k \log k)/10 < y$ be real numbers. Then we have*

$$\mathcal{M}_y(k) \leq e^{O(k \log\log y/\log y)} \left(\frac{2k}{ey \log y}\right)^k.$$

Remark 3.3. A similar bound was obtained by Bober, Goldmakher, Granville and Koukoulopoulos in the case of Steinhaus Random multiplicative functions (see the end of Section 5 of [Bober et al. 2018]). However, our argument is easier, and can be adapted to recover this case as well.

Proof of Proposition 3.2. By expanding the moment and using that

$$\mathbb{E}(\mathbb{X}(n)) = \begin{cases} 1 & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise,} \end{cases}$$

we get

$$\mathcal{M}_y(k) = \mathbb{E}\Big(\sum_{P^-(n)>y} \frac{\widetilde{d_{2k}}(n)}{n} \mathbb{X}(n)\Big) = \sum_{P^-(n)>y} \frac{\widetilde{d_{2k}}(n^2)}{n^2}, \tag{3-6}$$

where

$$\widetilde{d_k}(n) := \big|\big\{(n_1, \dots, n_k) \in \mathbb{N}^k, \text{ such that } n_j > 1 \text{ for all } j, \text{ and } n_1 \cdots n_k = n\big\}\big|.$$

Therefore, it follows from (3-3) that

$$\mathcal{M}_y(k) \leq \sum_{P^-(n)>y} \frac{d_{2k}(n^2)}{n^2} = \exp\Big(O\Big(\frac{k^2}{y \log y}\Big)\Big). \tag{3-7}$$

To obtain a better bound for $\mathcal{M}_y(k)$ we consider the following "good" event

$$\mathcal{A} := \{|\mathbb{Y}| \leq \varepsilon\}, \quad \text{where } \mathbb{Y} := \sum_{p>y} \frac{\mathbb{X}(p)}{p},$$

and $0 < \varepsilon < 1$ is a small parameter to be chosen. Note that

$$\mathbb{P}(\mathcal{A}^c) \leq \varepsilon^{-2\ell} \mathbb{E}(|\mathbb{Y}|^{2\ell}) \ll \left(\frac{3\ell}{e\varepsilon^2 y \log y}\right)^\ell,$$

since

$$\mathbb{E}(|\mathbb{Y}|^{2\ell}) = \sum_{\substack{p_1, \dots, p_{2\ell} > y \\ p_1 \cdots p_{2\ell} = \square}} \frac{1}{p_1 \dots p_{2\ell}} \leq \frac{(2\ell)!}{2^\ell \ell!} \Big(\sum_{p>y} \frac{1}{p^2}\Big)^\ell \ll e^{O(\ell/\log y)} \left(\frac{2\ell}{ey \log y}\right)^\ell, \tag{3-8}$$

which follows from the bounds $(2\ell)!/(2^\ell \ell!) \ll (2\ell/e)^\ell$ by Stirling's formula, and

$$\sum_{p>y} 1/p^2 = 1/(y \log y) + O\big(1/(y(\log y)^2)\big)$$

by the prime number theorem. Choosing $\ell = \lfloor (\varepsilon^2 y \log y)/3 \rfloor$ gives

$$\mathbb{P}(\mathcal{A}^c) \ll \exp\Big(-\frac{\varepsilon^2 y \log y}{3}\Big). \tag{3-9}$$

We now split the moment $\mathcal{M}_y(k)$ into two parts

$$\mathcal{M}_y(k) = \mathbb{E}\Big(\Big|\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n}\Big|^{2k} \cdot \mathbf{1}_{\mathcal{A}}\Big) + \mathbb{E}\Big(\Big|\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n}\Big|^{2k} \cdot \mathbf{1}_{\mathcal{A}^c}\Big), \tag{3-10}$$

where $\mathbf{1}_{\mathcal{B}}$ is the indicator function of the event $\mathcal{B}$. By the Cauchy–Schwarz inequality and the estimates (3-7) and (3-9), the contribution of the second part is

$$\mathbb{E}\Big(\Big|\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n}\Big|^{2k} \cdot \mathbf{1}_{\mathcal{A}^c}\Big) \le \mathcal{M}_y(2k)^{1/2} \cdot \mathbb{P}(\mathcal{A}^c)^{1/2} \ll \exp\Big(-\frac{\varepsilon^2 y\log y}{6} + O\Big(\frac{k^2}{y\log y}\Big)\Big). \tag{3-11}$$

Next, we shall estimate the contribution of the first part in (3-10). Note that on the event $\mathcal{A}$ we have $|e^{\mathbb{Y}} - 1| \le e^{\varepsilon}|\mathbb{Y}|$ and

$$\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n} = -1 + \prod_{p>y}\Big(1 - \frac{\mathbb{X}(p)}{p}\Big)^{-1} = -1 + e^{\mathbb{Y} + O(1/y\log y)} = e^{\mathbb{Y}} - 1 + O\Big(\frac{1}{y\log y}\Big).$$

Therefore, using Minkowski's inequality and (3-8) we derive

$$\mathbb{E}\Big(\Big|\sum_{\substack{n>1\\ P^-(n)>y}} \frac{\mathbb{X}(n)}{n}\Big|^{2k} \cdot \mathbf{1}_{\mathcal{A}}\Big)^{1/2k} \le e^{\varepsilon}\mathbb{E}(|\mathbb{Y}|^{2k})^{1/2k} + O\Big(\frac{1}{y\log y}\Big) \le e^{\varepsilon + O(1/\log y)}\sqrt{\frac{2k}{ey\log y}}.$$

Choosing $\varepsilon = (\log\log y)/\log y$ and combining this estimate with (3-11) completes the proof. □

4. The distribution of the tail in Pólya's Fourier expansion: Proof of Theorem 2.1

We start by recording two large sieve inequalities for quadratic characters, the most important of which is due to Heath-Brown [1995].

Lemma 4.1. *Let $x, N \ge 2$. Then for arbitrary complex numbers a_n we have*

$$\sum_{d\in\mathcal{F}(x)}\Big|\sum_{n\le N} a_n\chi_d(n)\Big|^2 \ll (x + N^2\log N)\sum_{\substack{m,n\le N\\ mn=\square}}|a_m a_n|, \tag{4-1}$$

and for any $\varepsilon > 0$ we have

$$\sum_{d\in\mathcal{F}(x)}\Big|\sum_{n\le N} a_n\chi_d(n)\Big|^2 \ll_{\varepsilon} (xN)^{\varepsilon}(x+N)\sum_{\substack{m,n\le N\\ mn=\square}}|a_m a_n|. \tag{4-2}$$

Proof. The first inequality is standard and is a straightforward application of the Pólya–Vinogradov inequality. It can be found for example in Lemma 1 of [Baker and Montgomery 1990], and can be traced back to [Elliott 1970]. The second inequality, which is deeper, was established by Heath-Brown [1995, Corollary 2]. □

Let $Y := \exp(C \log_2 x \log_3 x / \log_4 x)$, where $C > 0$ is a suitably large constant, and $Y \le N \le x^{21/40}$ be a real number. Using Lemma 4.1, we shall first prove that for all fundamental discriminants $|d| \le x$ except for a small exceptional set $\mathcal{E}(x)$, the quantity

$$\max_{\alpha\in[0,1)} \left| \sum_{N\le n\le 2N} \frac{\chi_d(n) a_n e(n\alpha)}{n} \right| \tag{4-3}$$

is small, where $\{a_n\}_{n\ge1}$ is an arbitrary sequence of complex numbers such that $|a_n| \le 1$ for all n. We shall use Heath-Brown's large sieve (4-2) if N is in the range $x^\varepsilon \le N \le x^{21/40}$, and Elliott's large sieve (4-1) in the remaining range $Y \le N \le x^\varepsilon$.

Proposition 4.2. *Let $A, \varepsilon > 0$ be fixed and put $\delta = 1/100$. Let $\{a_n\}_{n\ge1}$ be an arbitrary sequence of complex numbers such that $|a_n| \le 1$ for all n. Let N_1, N_2 be real numbers such that $x^\varepsilon \le N_1 < N_2 \le 2N_1 \le x^{21/40}$ be real numbers. Then there are at most $O_{\varepsilon,A}(x^{1-\delta})$ fundamental discriminants $|d| \le x$ such that*

$$\max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right| \ge \frac{1}{(\log N_1)^A}.$$

Proposition 4.3. *Let $A > 0$ be a fixed constant. Let $\{a_n\}_{n\ge1}$ be an arbitrary sequence of complex numbers such that $|a_n| \le 1$ for all n. There exist positive constants ε and C (which depend at most on A) such that for all real numbers N_1, N_2 verifying $\exp(C \log_2 x \log_3 x / \log_4 x) \le N_1 < N_2 \le 2N_1 \le x^\varepsilon$, the number of fundamental discriminants $|d| \le x$ such that*

$$\max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right| \ge \frac{1}{(\log N_1)^A}$$

is

$$\ll x \exp\left(-\frac{\log x \log_4 x}{10 \log_2 x} \right).$$

To prove these results we shall bound suitable moments of (4-3). We start with the following easy lemma which reduces the problem of bounding these moments to a setting where we can apply the large sieve.

Lemma 4.4. *Let $\{a_n\}_{n\ge1}$ be an arbitrary sequence of complex numbers such that $|a_n| \le 1$ for all n. Let $\mathcal{D}$ be a set of Dirichlet characters, and $2 \le N_1 < N_2 \le R$ be real numbers. Define $\mathcal{A} = \{b/R : 1 \le b \le R\}$. Then for any positive integer $k \ge 1$ we have*

$$\left(\sum_{\chi\in\mathcal{D}} \max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n) a_n e(\alpha n)}{n} \right|^{2k} \right)^{1/2k} \le \left(\sum_{\alpha\in\mathcal{A}} \sum_{\chi\in\mathcal{D}} \left| \sum_{N_1^k\le n\le N_2^k} \frac{\chi(n) g_{N_1,N_2,k}(n,\alpha)}{n} \right|^2 \right)^{1/2k} + O\big(|\mathcal{D}|^{1/2k} N_2 R^{-1}\big),$$

where

$$g_{N_1,N_2,k}(n,\alpha) := \sum_{\substack{N_1\le n_1,\dots,n_k\le N_2 \\ n_1\cdots n_k = n}} \prod_{j=1}^{k} a_{n_j} e(\alpha n_j), \tag{4-4}$$

and the implicit constant in the error term is absolute.

Proof. First, we observe that for all $\alpha \in [0,1)$ there exists $\beta_\alpha \in \mathcal{A}$ such that $|\alpha - \beta_\alpha| \le 1/R$. In this case we have $e(\alpha n) = e(\beta_\alpha n) + O(n/R)$, and hence

$$\max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\alpha n)}{n} \right| = \max_{\beta\in\mathcal{A}} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\beta n)}{n} \right| + O\left(\frac{N_2}{R}\right).$$

Therefore, it follows from Minkowski's inequality that

$$\begin{aligned}\Bigg(\sum_{\chi\in\mathcal{D}} \max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\alpha n)}{n} \right|^{2k} \Bigg)^{1/2k} &\\ \le \Bigg(\sum_{\chi\in\mathcal{D}} \max_{\alpha\in\mathcal{A}} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\alpha n)}{n} \right|^{2k} \Bigg)^{1/2k} &+ O\big(|\mathcal{D}|^{1/2k} N_2 R^{-1}\big)\\ \le \Bigg(\sum_{\alpha\in\mathcal{A}} \sum_{\chi\in\mathcal{D}} \left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\alpha n)}{n} \right|^{2k} \Bigg)^{1/2k} &+ O\big(|\mathcal{D}|^{1/2k} N_2 R^{-1}\big). \qquad (4\text{-}5)\end{aligned}$$

The lemma follows upon noting that

$$\left| \sum_{N_1\le n\le N_2} \frac{\chi(n)a_n e(\alpha n)}{n} \right|^{2k} = \left| \sum_{N_1^k\le n\le N_2^k} \frac{\chi(n) g_{N_1,N_2,k}(n,\alpha)}{n} \right|^2. \qquad \square$$

Proof of Proposition 4.2. Let

$$k := \begin{cases} 2 & \text{if } \sqrt{x} \le N_1 \le x^{21/40}, \\ \lfloor \log x / \log N_1 \rfloor & \text{if } x^{\varepsilon} \le N_1 < \sqrt{x}. \end{cases}$$

We observe that $2 \le k \le 1/\varepsilon$ by our assumption on N_1, and that $|g_{N_1,N_2,k}(n,\alpha)| \le d_k(n)$ for all N_1, N_2 and α. Let $\delta = 1/100$ and $\nu > 0$ be a small parameter to be chosen. Using Lemma 4.4 with the choice $R = N_1^{1+\delta}$ together with the large sieve inequality (4-2) and the easy inequality $|a+b|^k \le 2^k(|a|^k + |b|^k)$ we obtain

$$\begin{aligned}\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \left| \sum_{N_1\le n\le N_2} \frac{\chi_d(n)a_n e(\alpha n)}{n} \right|^{2k} &\\ \ll_\varepsilon \sum_{\alpha\in\mathcal{A}} \sum_{d\in\mathcal{F}(x)} \left| \sum_{N_1^k\le n\le N_2^k} \frac{\chi_d(n) g_{N_1,N_2,k}(n,\alpha)}{n} \right|^2 &+ \frac{x}{N_1^{2k\delta}}\\ \ll_{\varepsilon,\nu} (N_2^k x)^{\nu}(x+N_2^k) \sum_{\alpha\in\mathcal{A}} \sum_{\substack{N_1^k\le n_1,n_2\le N_2^k \\ n_1n_2=\square}} \frac{|g_{N_1,N_2,k}(n_1,\alpha) g_{N_1,N_2,k}(n_2,\alpha)|}{n_1n_2} &+ \frac{x}{N_1^{2k\delta}}\\ \ll_{\varepsilon,\nu} x^{21/20+3\nu} N_1^{1+\delta} \sum_{\substack{N_1^k\le n_1,n_2\le N_2^k \\ n_1n_2=\square}} \frac{d_k(n_1)d_k(n_2)}{n_1n_2} &+ \frac{x}{N_1^{2k\delta}}, \qquad (4\text{-}6)\end{aligned}$$

since $N_2^k \ll_\varepsilon x^{21/20}$ by our assumption on k and N_1. Now, writing $n_1n_2 = n^2$, we get

$$\sum_{\substack{N_1^k \le n_1, n_2 \le N_2^k \\ n_1n_2 = \square}} \frac{d_k(n_1)d_k(n_2)}{n_1n_2} \le \sum_{n \ge N_1^k} \frac{d_{2k}(n^2)}{n^2} \le N_1^{-k(1-\nu)} \sum_{n=1}^{\infty} \frac{d_{2k}(n^2)}{n^{1+\nu}} \ll_{\varepsilon,\nu} N_1^{-k(1-\nu)}. \tag{4-7}$$

Now, since $k \ge 2$, then $N_1^{3k/2} \ge N_1^{k+1} \ge x$, by our definition of k, which implies that $N_1^k \ge x^{2/3}$. Thus, choosing ν to be suitably small, and combining this estimate with (4-6) gives

$$\sum_{d \in \mathcal{F}(x)} \max_{\alpha \in [0,1)} \left| \sum_{N_1 \le n \le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right|^{2k} \ll_{\varepsilon,\nu} x^{21/20+3\nu} N_1^{1+\delta} x^{-2/3(1-\nu)} + \frac{x}{N_1^{2k\delta}} \ll_\varepsilon x^{1-4\delta/3}.$$

Finally, the number of fundamental discriminants $d \in \mathcal{F}(x)$ such that

$$\max_{\alpha \in [0,1)} \left| \sum_{N_1 \le n \le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right| \ge \frac{1}{(\log N_1)^A}$$

is

$$\le (\log N_1)^{2Ak} \sum_{d \in \mathcal{F}(x)} \max_{\alpha \in [0,1)} \left| \sum_{N_1 \le n \le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right|^{2k} \ll_{\varepsilon,A} x^{1-\delta},$$

which completes the proof. □

Proof of Proposition 4.3. We proceed similarly to the proof of Proposition 4.2 but we choose $k = \lfloor \log x/(3 \log N_1) \rfloor$ in this case. Then we have

$$1/(3\varepsilon) - 1 \le k \le (\log x \log_4 x)/(3C \log_2 x \log_3 x)$$

by our assumption on N_1. Using Lemma 4.4 with $R = N_1^2$ together with the large sieve inequality (4-1) and the easy inequality $|a+b|^k \le 2^k(|a|^k + |b|^k)$ we obtain

$$\begin{aligned}
\sum_{d \in \mathcal{F}(x)} \max_{\alpha \in [0,1)} \left| \sum_{N_1 \le n \le N_2} \frac{\chi_d(n) a_n e(\alpha n)}{n} \right|^{2k} &\ll e^{O(k)} \sum_{\alpha \in \mathcal{A}} \sum_{d \in \mathcal{F}(x)} \left| \sum_{N_1^k \le n \le N_2^k} \frac{\chi_d(n) g_{N_1,N_2,k}(n,\alpha)}{n} \right|^2 + \frac{x e^{O(k)}}{N_1^{2k}} \\
&\ll x e^{O(k)} \sum_{\alpha \in \mathcal{A}} \sum_{\substack{N_1^k \le n_1, n_2 \le N_2^k \\ n_1n_2 = \square}} \frac{|g_{N_1,N_2,k}(n_1,\alpha) g_{N_1,N_2,k}(n_2,\alpha)|}{n_1n_2} + \frac{x e^{O(k)}}{N_1^{2k}} \\
&\ll x N_1^2 e^{O(k)} \sum_{n \ge N_1^k} \frac{d_{2k}(n^2)}{n^2} + \frac{x}{N_1^{2k}}
\end{aligned} \tag{4-8}$$

by (4-7), since $|g_{N_1,N_2,k}(n,\alpha)| \le d_k(n)$ for all N_1, N_2 and α. To bound the sum over n, we shall use Rankin's trick. Choosing $\nu = \log_3 k/(2\log k)$ and using (3-2) we obtain

$$\sum_{n\ge N_1^k} \frac{d_{2k}(n^2)}{n^2} \le N_1^{-k\nu} \sum_{n=1}^{\infty} \frac{d_{2k}(n^2)}{n^{2-\nu}} \ll \exp\Big(-\frac{k\log_3 k\log N_1}{2\log k} + O(k\log_2 k)\Big).$$

Inserting this estimate in (4-8) we deduce that

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big|\sum_{N_1\le n\le N_2} \frac{\chi_d(n)a_n e(\alpha n)}{n}\Big|^{2k} \ll x\exp\Big(-\frac{k\log_3 k\log N_1}{2\log k} + O(k\log_2 k + \log N_1)\Big).$$

Therefore, the number of fundamental discriminants $d \in \mathcal{F}(x)$ such that

$$\max_{\alpha\in[0,1)} \Big|\sum_{N_1\le n\le N_2} \frac{\chi_d(n)a_n e(\alpha n)}{n}\Big| \ge \frac{1}{(\log N_1)^A}$$

is

$$\begin{aligned}
&\le (\log N_1)^{2Ak} \sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big|\sum_{N_1\le n\le N_2} \frac{\chi_d(n)a_n e(\alpha n)}{n}\Big|^{2k}\\
&\ll x\exp\Big(-\frac{k\log_3 k\log N_1}{2\log k} + O(k\log_2 k + Ak\log_2 N_1 + \log N_1)\Big)\\
&\ll x\exp\Big(-\frac{\log x\log_4 x}{10\log_2 x}\Big)
\end{aligned}$$

by our assumption on N_1 and k, if ε is suitably small and C is suitably large. This completes the proof. □

Using Propositions 4.2 and 4.3 reduces the proof of Theorem 2.1 to studying the distribution of the maximum over $\alpha \in [0, 1)$ of the very short sum

$$\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|.$$

To this end we shall use the large sieve to bound its $2k$-th moments for *every* $k \le (\log x)/(3\log Y)$.

Proposition 4.5. *Let $h(n)$ be a completely multiplicative function such that $|h(n)| \le 1$ for all n. Let $C > 0$ be a suitably large constant and put $Y = \exp(C\log_2 x\log_3 x/\log_4 x)$. For any positive integer $2 \le k \le \log x/(3\log Y)$ and real number $(k\log k)/10 \le y \le k^2$ we have*

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k} \ll x\Big(e^{O\left(\frac{k\log_2 k}{\log k}\right)}\Big(\frac{2e^{2\gamma}k\log y}{ey}\Big)^k + O\Big(\frac{1}{(\log k)^{10k}}\Big)\Big).$$

Proof of Proposition 4.5. We first define

$$S_d(y,Y)=\max_{\alpha\in[0,1)}\Big|\sum_{\substack{2\le n\le Y\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|. \tag{4-9}$$

Then we have

$$\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}=\sum_{\substack{1\le a\le Y\\ P^+(a)\le y}}\frac{\chi_d(a)h(a)}{a}\sum_{\substack{1<b\le Y/a\\ P^-(b)>y}}\frac{\chi_d(b)h(b)e(\alpha ab)}{b},$$

and hence

$$\max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|\le\sum_{\substack{1\le a\le Y\\ P^+(a)\le y}}\frac{S_d(y,Y/a)}{a}.$$

Therefore, using Hölder's inequality we obtain

$$\begin{aligned}\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}&\le\sum_{d\in\mathcal{F}(x)}\Big(\sum_{P^+(a)\le y}\frac{1}{a}\Big)^{2k-1}\sum_{\substack{1\le a\le Y\\ P^+(a)\le y}}\frac{S_d(y,Y/a)^{2k}}{a}\\&\le\big(e^{\gamma}\log y+O(1)\big)^{2k-1}\sum_{\substack{1\le a\le Y\\ P^+(a)\le y}}\frac{1}{a}\sum_{d\in\mathcal{F}(x)}S_d(y,Y/a)^{2k}.\end{aligned} \tag{4-10}$$

by Mertens' theorem. We shall now bound the moments

$$\sum_{d\in\mathcal{F}(x)}S_d(y,z)^{2k}$$

uniformly in $2\le y<z\le Y$. To this end we split the inner sum over n into two parts $y<n\le N$, and $N<n\le z$, where

$$N=\min\big(z,\exp(C\log k\log_2 k/\log_3 k)\big). \tag{4-11}$$

Note that the second part will be empty unless $z>\exp(C\log k\log_2 k/\log_3 k)$. Using Minkowski's inequality we get

$$\begin{aligned}\Big(\sum_{d\in\mathcal{F}(x)}S_d(y,z)^{2k}\Big)^{1/2k}&\le\Big(\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{2\le n\le N\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}\Big)^{1/2k}\\&\quad+\Big(\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{N<n\le z\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}\Big)^{1/2k}.\end{aligned} \tag{4-12}$$

We start by bounding the first term. By Lemma 4.4 with $R:=\exp(2C\log k\log_2 k/\log_3 k)$ and

$$a_n=\begin{cases}h(n) & \text{if } P^-(n)>y\\ 0 & \text{otherwise,}\end{cases}$$

we obtain

$$\Big(\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{2\le n\le N\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}\Big)^{1/2k}$$
$$\le\Big(\sum_{\alpha\in\mathcal{A}}\sum_{d\in\mathcal{F}(x)}\Big|\sum_{\substack{n\le N^k\\ P^-(n)>y}}\frac{\chi_d(n)g_{2,N,k}(n,\alpha)}{n}\Big|^2\Big)^{1/2k}+O\big(x^{1/2k}R^{-1/2}\big),\quad(4\text{-}13)$$

where $\mathcal{A}=\{b/R:1\le b\le R\}$. We now proceed similarly to the proof of (4-8). Using the large sieve inequality (4-1), and noting that $N^k\le x^{1/3}$ and $|g_{2,N,k}(n,\alpha)|\le\widetilde{d}_k(n)$ for all α and N we derive

$$\sum_{d\in\mathcal{F}(x)}\Big|\sum_{\substack{n\le N^k\\ P^-(n)>y}}\frac{\chi_d(n)g_{2,N,k}(n,\alpha)}{n}\Big|^2\ll x\sum_{\substack{n_1,n_2\le N^k\\ P^-(n_1n_2)>y\\ n_1n_2=\square}}\frac{\widetilde{d}_k(n_1)\widetilde{d}_k(n_2)}{n_1n_2}\le x\sum_{P^-(n)>y}\frac{\widetilde{d_{2k}}(n^2)}{n^2}$$
$$\ll xe^{O(k\log\log y/\log y)}\Big(\frac{2k}{ey\log y}\Big)^k,\quad(4\text{-}14)$$

where the last inequality follows from (3-6) and Proposition 3.2. Inserting this estimate in (4-13) implies that

$$\Big(\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{2\le n\le N\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}\Big)^{1/2k}\le x^{1/(2k)}\Big(e^{O(\log_2 k/\log k)}\Big(\frac{2k}{ey\log y}\Big)^{1/2}+O\Big(\frac{1}{R^{1/2}}\Big)\Big).\quad(4\text{-}15)$$

We now assume that $z>\exp(C\log k\log_2 k/\log_3 k)$ and bound the second term on the right-hand side of (4-12). We shall split the inner sum over n into dyadic intervals. Let $J_1=\lfloor\log N/\log 2\rfloor$, $J_2=\lfloor\log z/\log 2\rfloor$, and define $t_{J_1}=N$, $t_{J_2+1}=z$, and $t_j:=2^j$ for $J_1+1\le j\le J_2$. Using Hölder's inequality we obtain

$$\Big|\sum_{\substack{N<n\le z\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}=\Big|\sum_{J_1\le j\le J_2}\frac{1}{j^2}\Big(j^2\sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big)\Big|^{2k}$$
$$\le\Big(\sum_{J_1\le j\le J_2}\frac{1}{j^{\frac{4k}{2k-1}}}\Big)^{2k-1}\sum_{J_1\le j\le J_2}j^{4k}\Big|\sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}$$
$$\le\Big(\frac{C_1}{\log N}\Big)^{2k+1}\sum_{J_1\le j\le J_2}j^{4k}\Big|\sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k},\quad(4\text{-}16)$$

for some constant $C_1 > 0$. Therefore, this reduces the problem to bounding the following moments over dyadic intervals $[t_j, t_{j+1}]$:

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}.$$

Let $\mathcal{B}_j = \{b/4^j : 1 \le b \le 4^j\}$. By the easy inequality $|a+b|^k \le 2^k(|a|^k + |b|^k)$ together with Lemma 4.4 with $R = 4^j$ and the same choice of a_n as before, we derive

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k} \ll e^{O(k)} \sum_{\alpha\in\mathcal{B}_j}\sum_{d\in\mathcal{F}(x)} \Big| \sum_{\substack{t_j^k<n\le t_{j+1}^k\\ P^-(n)>y}} \frac{\chi_d(n)g_{t_j,t_{j+1},k}(n,\alpha)}{n}\Big|^2 + \frac{xe^{O(k)}}{2^{2jk}}, \quad (4\text{-}17)$$

since $t_j \asymp t_{j+1} \asymp 2^j$. Furthermore, similarly to the proof of (4-8), it follows from the large sieve inequality (4-1) that the main term on the right-hand side of (4-17) is

$$\ll xe^{O(k)} \sum_{\alpha\in\mathcal{B}_j} \sum_{\substack{t_j^k<n_1,n_2\le t_{j+1}^k\\ P^-(n_1n_2)>y\\ n_1n_2=\square}} \frac{|g_{t_j,t_{j+1},k}(n_1,\alpha)g_{t_j,t_{j+1},k}(n_2,\alpha)|}{n_1n_2} \ll e^{O(k)}4^j x \sum_{n\ge t_j^k} \frac{d_{2k}(n^2)}{n^2}, \quad (4\text{-}18)$$

since $t_{j+1}^k \le z^k \le Y^k \le x^{1/3}$. We now put $\nu = \log_3 k/(2\log k)$ and use (3-2) to get

$$\sum_{n\ge t_j^k} \frac{d_{2k}(n^2)}{n^2} \le t_j^{-k\nu} \sum_{n=1}^{\infty} \frac{d_{2k}(n^2)}{n^{2-\nu}} \ll \exp\Big(-\frac{jk\log_3 k}{4\log k} + O(k\log_2 k)\Big). \quad (4\text{-}19)$$

Inserting these estimates in (4-17) gives

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k} \ll x\exp\Big(-\frac{jk\log_3 k}{5\log k} + O(k\log_2 k)\Big).$$

Using this last estimate together with (4-16), and noting that $j^4 \le \exp(j\log_3 k/(20\log k))$ for all $j \ge J_1$ by our choice of N if C is suitably large, we derive

$$\begin{aligned}\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{N<n\le z\\ P^-(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k} &\ll xe^{O(k\log_2 k)} \sum_{J_1\le j\le J_2} \exp\Big(-\frac{jk\log_3 k}{20\log k}\Big)\\ &\ll x\exp\Big(-\frac{(\log N)k\log_3 k}{20\log k} + O(k\log_2 k)\Big)\\ &\ll x\exp\Big(-\frac{C}{40}k\log_2 k\Big).\end{aligned}$$

Combining this estimate with (4-12) and (4-15) we obtain

$$\sum_{d\in\mathcal{F}(x)} S_d(y,z)^{2k} \ll x\bigg(e^{O(\log_2 k/\log k)}\bigg(\frac{2k}{ey\log y}\bigg)^{1/2} + O\bigg(\frac{1}{(\log k)^{C/80}}\bigg)\bigg)^{2k},$$

uniformly for $y < z \le Y$. Therefore, inserting this estimate in (4-10) we deduce

$$\sum_{d\in\mathcal{F}(x)} \max_{\alpha\in[0,1)} \bigg|\sum_{\substack{n\le Y\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg|^{2k} \ll x\bigg(e^{O\left(\frac{\log_2 k}{\log k}\right)}\bigg(\frac{2e^{2\gamma}k\log y}{ey}\bigg)^{1/2} + O\bigg(\frac{1}{(\log k)^{C/80-1}}\bigg)\bigg)^{2k}.$$

The result follows upon choosing C to be suitably large, and using the basic inequality[6] $(a+b)^{2m} \le (a^{2m}+\sqrt{b}^{2m})(1+\sqrt{b})^{2m} \le (a^{2m}+b^m)e^{2m\sqrt{b}}$, which is valid for all real numbers $a, b > 0$ and positive integers m. □

We end this section by deducing Theorem 2.1 from Propositions 4.2, 4.3, and 4.5.

Proof of Theorem 2.1. Let $Y := \exp(C\log_2 x\log_3 x/\log_4 x)$ for some suitably large constant $C > 0$. Let $L_1 := \lfloor \log Y/\log 2\rfloor$, $L_2 := \lfloor \log Z/\log 2\rfloor$, and define $s_{L_1} := Y$, $s_{L_2+1} := Z$, and $s_\ell := 2^\ell$ for $L_1+1 \le \ell \le L_2$. Then we have

$$\begin{aligned}\max_{\alpha\in[0,1)} \bigg|\sum_{\substack{1\le n\le Z\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg| &\le \max_{\alpha\in[0,1)} \bigg|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg| + \sum_{L_1\le\ell\le L_2} \max_{\alpha\in[0,1)} \bigg|\sum_{\substack{s_\ell\le n\le s_{\ell+1}\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg|. \end{aligned} \tag{4-20}$$

Using Propositions 4.2 and 4.3 with $A = 2$ and

$$a_n = \begin{cases} h(n) & \text{if } P^+(n) > y,\\ 0 & \text{otherwise,}\end{cases}$$

we deduce that

$$\sum_{L_1\le\ell\le L_2} \max_{\alpha\in[0,1)} \bigg|\sum_{\substack{s_\ell\le n\le s_{\ell+1}\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg| \ll \sum_{L_1\le\ell\le L_2} \frac{1}{\ell^2} \ll \frac{\log_4 x}{\log_2 x\log_3 x}, \tag{4-21}$$

for all fundamental discriminants $|d| \le x$ except for a set $\mathcal{E}(x)$ of size

$$|\mathcal{E}(x)| \ll L_2 x\exp\bigg(-\frac{\log x\log_4 x}{10\log_2 x}\bigg) \ll x\exp\bigg(-\frac{\log x\log_4 x}{20\log_2 x}\bigg).$$

For $B > 0$ we let $\widetilde{\Psi}_{x,y}(B)$ be the proportion of fundamental discriminants $|d| \le x$ such that

$$\max_{\alpha\in[0,1)} \bigg|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}} \frac{\chi_d(n)h(n)e(\alpha n)}{n}\bigg| > e^\gamma B.$$

[6]This inequality simply follows by considering the two cases $a \le \sqrt{b}$ and $a \ge \sqrt{b}$.

Let k be a positive integer satisfying the assumptions of Proposition 4.5. Then, it follows from this result that

$$\widetilde{\Psi}_{x,y}(B) \le (e^{\gamma}B)^{-2k}\frac{1}{|\mathcal{F}(x)|}\sum_{d\in\mathcal{F}(x)}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\chi_d(n)h(n)e(\alpha n)}{n}\Big|^{2k}$$
$$\ll e^{O\left(\frac{k\log_2 k}{\log k}\right)}\left(\frac{2k\log y}{eB^2y}\right)^k + O\left(\frac{1}{(e^{\gamma}B(\log k)^5)^{2k}}\right).$$

We now assume that $1/(\log y)^2 \le B \le \sqrt{20}$ and choose $k = \lfloor (B^2y)/(2\log y)\rfloor$. This gives

$$\widetilde{\Psi}_{x,y}(B) \ll \exp\left(-\frac{B^2y}{2\log y}\left(1+O\left(\frac{\log_2 y}{\log y}\right)\right)\right). \tag{4-22}$$

Combining this estimate with (4-20) and (4-21) and choosing $B = A - C_0\log_4 x/(\log_2 x\log_3 x)$ for some suitably large constant C_0 completes the proof. □

5. The distribution of the tail of the Pólya Fourier series for prime discriminants: Proof of Theorem 2.7

In order to prove Theorem 2.7 we follow the same lines of the proof of Theorem 2.1 but we replace the large sieve inequality (4-1) by the following large sieve inequality for prime discriminants, which is a special case of Lemma 9 of [Montgomery and Vaughan 1979] (see also Lemma 1 of [Baker and Montgomery 1990]).

Lemma 5.1 [Montgomery and Vaughan 1979, Lemma 9]. *Let x, N be real numbers such that $x \ge 2$ and $2 \le N \le x^{1/3}$. Then for arbitrary complex numbers $a_1, \dots, a_N$ we have*

$$\sum_{p\le x}\Big|\sum_{n\le N}a_n\psi_p(n)\Big|^2 \ll \frac{x}{\log x}\sum_{\substack{m,n\le N\\ mn=\square}}|a_ma_n|.$$

Since the number of primes up to x is smaller by a factor of size $\log x$ compared to the number of fundamental discriminants up to x, it suffices to establish the analogue of Proposition 4.5 for prime discriminants. Indeed, the savings in Propositions 4.2 and 4.3 are much larger than $\log x$, and hence we can use these results in this setting as well by simply embedding the set of primes $p \le x$ in the set of fundamental discriminants d with $|d| \le x$.

Proposition 5.2. *Let $h(n)$ be a completely multiplicative function such that $|h(n)| \le 1$ for all n. Let $Y = \exp(C\log_2 x\log_3 x/\log_4 x)$, for some suitably large constant C. Then, for any positive integer $2 \le k \le \log x/(3\log Y)$ and real number $(k\log k)/10 \le y \le k^2$ we have*

$$\sum_{p\le x}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big|^{2k} \ll \pi(x)\left(e^{O\left(\frac{k\log_2 k}{\log k}\right)}\left(\frac{2e^{2\gamma}k\log y}{ey}\right)^k + O\left(\frac{1}{(\log k)^{10k}}\right)\right).$$

Proof. We shall closely follow the proof of Proposition 4.5 and only indicate where the main changes occur. First we define

$$\widetilde{S}_p(y, Y) = \max_{\alpha\in[0,1)} \Big| \sum_{\substack{2\le n\le Y\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|. \tag{5-1}$$

Then, similarly to (4-10) we have

$$\sum_{p\le x} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{n\le Y\\ P^+(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k} \le \big(e^\gamma \log y + O(1)\big)^{2k-1} \sum_{\substack{a\le Y\\ P^+(a)\le y}} \frac{1}{a} \sum_{p\le x} \widetilde{S}_p(y, Y/a)^{2k}. \tag{5-2}$$

Let $y < z \le Y$ be a real number and N be defined by (4-11). Using Minkowski's inequality as in (4-12) we get

$$\begin{aligned}\Big(\sum_{p\le x} \widetilde{S}_p(y, z)^{2k}\Big)^{1/2k} &\le \Big(\sum_{p\le x} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{2\le n\le N\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k}\Big)^{1/2k}\\ &\quad + \Big(\sum_{p\le x} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{N<n\le z\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k}\Big)^{1/2k}. \end{aligned}\tag{5-3}$$

We start by bounding the first term. Let $R = \exp(2C \log k \log_2 k/\log_3 k)$. Using the same argument leading to (4-15) and replacing the large sieve inequality (4-1) by Lemma 5.1 we obtain

$$\begin{aligned}&\Big(\sum_{p\le x} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{2\le n\le N\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k}\Big)^{1/2k}\\ &\qquad \le \pi(x)^{1/(2k)} \bigg(e^{O(\log_2 k/\log k)} \Big(\frac{2k}{ey\log y}\Big)^{1/2} + O\Big(\frac{1}{R^{1/2}}\Big)\bigg). \end{aligned}\tag{5-4}$$

We now assume that $z > \exp(C \log k \log_2 k/\log_3 k)$ and bound the second term on the right-hand side of (5-3). By (4-16) we have

$$\Big| \sum_{\substack{N<n\le z\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k} \le \Big(\frac{C_1}{\log N}\Big)^{2k+1} \sum_{J_1\le j\le J_2} j^{4k} \Big| \sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k}, \tag{5-5}$$

for some constant $C_1 > 0$, where $J_1 = \lfloor \log N/\log 2\rfloor$, $J_2 = \lfloor \log z/\log 2\rfloor$, $t_{J_1} = N$, $t_{J_2+1} = z$, and $t_j = 2^j$ for $J_1 + 1 \le j \le J_2$. As before we let $\mathcal{B}_j = \{b/4^j : 1 \le b \le 4^j\}$. Combining Lemmas 4.4 and 5.1 with (4-18) and (4-19) we deduce that

$$\begin{aligned}\sum_{p\le x} \max_{\alpha\in[0,1)} \Big| \sum_{\substack{t_j<n\le t_{j+1}\\ P^-(n)>y}} \frac{\psi_p(n)h(n)e(\alpha n)}{n} \Big|^{2k} &\ll e^{O(k)} \sum_{\alpha\in\mathcal{B}_j} \sum_{p\le x} \Big| \sum_{\substack{t_j^k<n\le t_{j+1}^k\\ P^-(n)>y}} \frac{\psi_p(n) g_{t_j,t_{j+1},k}(n,\alpha)}{n} \Big|^2 + \frac{\pi(x)e^{O(k)}}{2^{2jk}}\\ &\ll \pi(x) \exp\Big(-\frac{jk\log_3 k}{5\log k} + O(k\log_2 k)\Big).\end{aligned}$$

Using this last estimate together with (5-5) we derive

$$\sum_{p\le x}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{N<n\le z\\ P^-(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big|^{2k}\ll \pi(x)\exp\Big(-\frac{C_1}{40}k\log_2 k\Big).$$

We have now established the key estimates over prime discriminants. Thus we continue the proof along the exact same lines as the proof of Proposition 4.5, by combining this last estimate with (5-2), (5-3) and (5-4). This yields the desired result. □

Proof of Theorem 2.7. As before we let $Y=\exp(C\log_2 x\log_3 x/\log_4 x)$ for some suitably large constant $C>0$, and put $L_1=\lfloor \log Y/\log 2\rfloor$, $L_2=\lfloor \log Z/\log 2\rfloor$, $s_{L_1}=Y$, $s_{L_2+1}=Z$, and $s_\ell=2^\ell$ for $L_1+1\le \ell\le L_2$. By (4-20) we have

$$\max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Z\\ P^+(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big| \le \max_{\alpha\in[0,1)}\Big|\sum_{\substack{1\le n\le Y\\ P^+(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big|+\sum_{L_1\le\ell\le L_2}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{s_\ell\le n\le s_{\ell+1}\\ P^+(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big|.$$

By embedding the set of primes $3\le p\le x$ into the set of fundamental discriminants $|d|\le x$ (since for any such prime, p or $-p$ is a fundamental discriminant), it follows from (4-21) that

$$\sum_{L_1\le\ell\le L_2}\max_{\alpha\in[0,1)}\Big|\sum_{\substack{s_\ell\le n\le s_{\ell+1}\\ P^+(n)>y}}\frac{\psi_p(n)h(n)e(\alpha n)}{n}\Big|\ll\frac{\log_4 x}{\log_2 x\log_3 x}, \tag{5-6}$$

for all primes $p\le x$ except for a set $\mathcal{E}_1(x)$ of size

$$|\mathcal{E}_1(x)|\ll x\exp\Big(-\frac{\log x\log_4 x}{20\log_2 x}\Big).$$

The result follows along the same exact lines of the proof of Theorem 2.1, by replacing Proposition 4.5 by Proposition 5.2. □

6. The distribution of $L(1,\psi_p)$: proof of Theorem 1.10

6A. ***Proof of the estimate (1-8).*** In the range $2\le\tau\le(\log_2 x)/2-2\log_3 x$, we shall use [Lamzouri 2017] where we proved an asymptotic formula for complex moments of $L(1,\psi_p)$ involving a secondary term coming from a possible Laudau–Siegel exceptional discriminant. Indeed, it follows from Theorem 1.2 of [Lamzouri 2017] that for all real numbers $2\le k\le\sqrt{\log x}/(\log_2 x)^2$ we have

$$\frac{2}{\mathrm{Li}(x)}\sum_{\substack{p\le x\\ p\equiv 3 \bmod 4}}L(1,\psi_p)^k=\mathbb{E}\big(L(1,\mathbb{X})^k\big)+E_1+O\Big(\exp\Big(-\frac{\sqrt{\log x}}{10\log\log x}\Big)\Big), \tag{6-1}$$

where

$$|E_1| \le \left|\mathbb{E}\big(\mathbb{X}(|d_1|)L(1,\mathbb{X})^k\big)\right|,$$

and where d_1 is the possible "exceptional" discriminant[7] with $|d_1| \le \exp(\sqrt{\log x})$, and

$$L(1,\mathbb{X}) := \prod_{q \text{ prime}} \left(1 - \frac{\mathbb{X}(q)}{q}\right)^{-1},$$

with the $\mathbb{X}(q)$ being I.I.D. random variables taking the values ± 1 with probability $1/2$. We cannot rule out that the term E_1 is of the same size as the main term, but we will prove that it will not heavily affect the size of the k-th moment of $L(1,\psi_p)$ if k is large. Indeed, we observe that

$$\frac{\mathbb{E}\big(\mathbb{X}(|d_1|)L(1,\mathbb{X})^k\big)}{\mathbb{E}\big(L(1,\mathbb{X})^k\big)} = \prod_{q|d_1} \frac{\big(1-\frac{1}{q}\big)^{-k} - \big(1+\frac{1}{q}\big)^{-k}}{\big(1-\frac{1}{q}\big)^{-k} + \big(1+\frac{1}{q}\big)^{-k}} = \prod_{q|d_1} \frac{1-\delta_q}{1+\delta_q}, \tag{6-2}$$

where $\delta_q = \big(1 - \frac{2}{q+1}\big)^k$. Let $\varepsilon > 0$ be a suitably small constant. Let q_1 be the largest prime factor of d_1. Since d_1 is square-free and $d_1 > (\log x)^{10/\varepsilon}$, if x is large enough, by Siegel's theorem, we must have $q_1 > (\log|d_1|)/2 > (5/\varepsilon)\log\log x$, since otherwise $|d_1| < \prod_{q \le (\log|d_1|)/2} q = |d_1|^{1/2+o(1)}$ by the prime number theorem, which is a contradiction. Inserting this estimate in (6-2) gives

$$0 < \frac{\mathbb{E}\big(\mathbb{X}(|d_1|)L(1,\mathbb{X})^k\big)}{\mathbb{E}\big(L(1,\mathbb{X})^k\big)} \le 1 - \delta_{q_1} = 1 - \left(1 - \frac{2}{q_1+1}\right)^k \le 1 - \exp\left(-\varepsilon\frac{k}{\log k}\right), \tag{6-3}$$

for all real numbers $2 \le k \le \sqrt{\log x}/(\log_2 x)^2$. On the other hand, it follows from Proposition 1.2 of [Lamzouri 2010] that

$$\mathbb{E}\big(L(1,\mathbb{X})^k\big) = \exp\left(k\log_2 k + k\gamma + \frac{k}{\log k}\left(B_0 - 1 + O\left(\frac{1}{\log k}\right)\right)\right). \tag{6-4}$$

Combining this estimate with (6-1) and (6-3) we obtain

$$\frac{2}{\pi(x)} \sum_{\substack{p \le x \\ p \equiv 3 \bmod 4}} L(1,\psi_p)^k = \exp\left(k\log_2 k + k\gamma + \frac{k}{\log k}\big(B_0 - 1 + O(\varepsilon)\big)\right). \tag{6-5}$$

Finally, the estimate for the distribution function (1-8) follows from the proof of Theorem 0.1 of [Lamzouri 2010] (which holds for general random models of this type), with the choice $\tau = \log k + B_0$. This completes the proof of (1-8) in the case $a = 3$. We also note that the case $a = 1$ is similar, since one can derive the same estimate as (6-1) over the primes $p \equiv 1 \bmod 4$ using the method of [Lamzouri 2017].

[7] By the Landau–Page Theorem (see Chapter 20 of [Davenport 2000]), there is at most one square-free integer d_1 such that $|d_1| \le \exp(\sqrt{\log x})$ and $L(s,\chi_{d_1})$ has a zero in the region $\mathrm{Re}(s) > 1 - c/\sqrt{\log x}$, for some positive constant c. If it exists, we refer to such d_1 as the exceptional discriminant in the range $|d_1| \le \exp(\sqrt{\log x})$.

We now show how to obtain the same estimate for the proportion of primes $p \equiv a \bmod 4$ such that $L(1, \psi_p\chi_{-3}) > (2e^\gamma/3)\tau$. First, by a slight adaptation of the proof of Theorem 1.2 of [Lamzouri 2017] we get

$$\frac{2}{\mathrm{Li}(x)} \sum_{\substack{p\le x\\ p\equiv a \bmod 4}} L(1, \psi_p\chi_{-3})^k = \mathbb{E}\big(L(1, \mathbb{X}\chi_{-3})^k\big) + E_2 + O\bigg(\exp\bigg(-\frac{\sqrt{\log x}}{10\log\log x}\bigg)\bigg), \tag{6-6}$$

where

$$|E_2| \le \big|\mathbb{E}\big(\mathbb{X}(|d_1|)L(1, \mathbb{X}\chi_{-3})^k\big)\big|,$$

and where d_1 is the exceptional discriminant (if it exists) with $|d_1| \le \exp(\sqrt{\log x})$, and

$$L(1, \mathbb{X}\chi_{-3}) := \prod_{q \text{ prime}} \bigg(1 - \frac{\mathbb{X}(q)\chi_{-3}(q)}{q}\bigg)^{-1}.$$

By the independence of the $\mathbb{X}(q)$'s we observe that

$$\mathbb{E}\big(L(1, \mathbb{X}\chi_{-3})^k\big) = \prod_{q\neq 3} \mathbb{E}\bigg(\bigg(1 - \frac{\mathbb{X}(q)}{q}\bigg)^{-k}\bigg) = \bigg(\frac{2}{3}\bigg)^k \bigg(\frac{2}{1+2^{-k}}\bigg)\mathbb{E}\big(L(1, \mathbb{X})^k\big).$$

Furthermore, if we fix a suitably small constant $\varepsilon > 0$, then a similar argument leading to (6-3) gives

$$|E_2| \le \bigg(1 - \exp\bigg(-\varepsilon\frac{k}{\log k}\bigg)\bigg)\mathbb{E}\big(L(1, \mathbb{X}\chi_{-3})^k\big), \tag{6-7}$$

if x is large enough. Thus by (6-4) we deduce that

$$\frac{2}{\pi(x)} \sum_{\substack{p\le x\\ p\equiv a \bmod 4}} \bigg(\frac{3}{2}L(1, \psi_p\chi_{-3})\bigg)^k = \exp\bigg(k\log_2 k + k\gamma + \frac{k}{\log k}\big(B_0 - 1 + O(\varepsilon)\big)\bigg).$$

Finally, the result follows from the proof of Theorem 0.1 of [Lamzouri 2010].

Remark 6.1. We observe that assuming GRH, Holmin, Jones, Kurlberg, McLeman and Petersen [Holmin et al. 2019] established the asymptotic formula (6-1) without the term E_1 in the larger range $k \le (\log x)/(50(\log\log x)^2)$. This justifies Remark 1.8.

6B. ***Proof of* (1-9).** In the range $(\log_2 x)/2 - 2\log_3 x \le \tau \le \log_2 x - \log_3 x - C_2$, we need to use a different argument since asymptotic formulas for very large moments of $L(1, \psi_p)$ are not known, due to the lack of strong unconditional bounds on character sums over primes. In this case, our strategy is to use Theorem 2.7 to truncate $L(1, \psi_p)$ over y-friable integers, and then control the behaviour of $\psi_p(q)$ over the primes $q \le y$. To this end we establish the following lemma which follows from Bombieri's proof [1987, Chapter 6] of Linnik's theorem.

Lemma 6.2. *Let $\{\varepsilon_q\}_{q \text{ prime}}$ be a sequence of ± 1, and $a \in \{1, 3\}$. Let x be large and $3 \le y \le c_0 \log x$ be a real number, where c_0 is a suitably small constant. Let $\mathcal{P}(x, y, a, \{\varepsilon_q\})$ be the set of primes $p \equiv a \bmod 4$ with $p \le x$ such that $\psi_p(q) = \varepsilon_q$ for all primes $q \le y$. Then we have*

$$|\mathcal{P}(x, y, a, \{\varepsilon_q\})| \gg \pi(x) \exp\left(-3y + O\left(\frac{y}{\log y}\right)\right).$$

Proof. First recall that if $p \equiv 3 \bmod 4$ then $\psi_p(2) = 1$ if $p \equiv 7 \bmod 8$, and equals -1 if $p \equiv 3 \bmod 8$. Similarly, if $p \equiv 1 \bmod 4$ then $\psi_p(2) = 1$ if $p \equiv 1 \bmod 8$, and $\psi_p(2) = -1$ if $p \equiv 5 \bmod 8$. Let $3 \le q \le y$ be a prime number. By the law of quadratic reciprocity, there exists a residue class $b_q \bmod q$ such that if $p \equiv b_q \bmod q$ then $\psi_p(q) = \varepsilon_q$. We now define $Q := 8 \prod_{3 \le q \le y} q$. Then, by the Chinese remainder theorem there exists a residue class $b \bmod Q$ such that if $p \equiv b \bmod Q$ then $p \equiv a \bmod 4$ and $\psi_p(q) = \varepsilon_q$ for all primes $q \le y$. By Bombieri's proof [1987, Chapter 6] of Linnik's theorem we deduce that there exists a small constant $c > 0$ such that if $x > Q^{1/c}$ then

$$|\mathcal{P}(x, y, a, \{\varepsilon_q\})| \ge \pi(x; Q, b) \gg \frac{\pi(x)}{Q^3}.$$

The result follows upon noting that $Q = \exp(y(1 + O(1/\log y)))$ by the prime number theorem. □

Proof of the lower bound (1-9). Let $Z = x^{21/40}$. By (2-9) we have

$$L(1, \psi_p) = \sum_{n \le Z} \frac{\psi_p(n)}{n} + O(1),$$

for all primes $p \le x$. Let c_0 be the constant in Lemma 6.2, and $3 \le y \le c_0 \log x$ be a parameter to be chosen. Define $\mathcal{P}(y)$ to be the set of primes $p \le x$ such that $p \equiv a \bmod 4$, $\psi_p(q) = 1$ for all primes $q \le y$ and

$$\left| \sum_{\substack{1 \le n \le Z \\ P^+(n) > y_1}} \frac{\psi_p(n)}{n} \right| \le e^{\gamma}, \tag{6-8}$$

where $y_1 := 10y \log y$. Combining Lemma 6.2 with Theorem 2.7 yields

$$|\mathcal{P}(y)| \gg \pi(x) \exp(-4y). \tag{6-9}$$

On the other hand, by (2-10) and (6-8) we deduce that for every prime $p \in \mathcal{P}(y)$ we have

$$\begin{aligned} L(1, \psi_p) &= \sum_{\substack{1 \le n \le Z \\ P^+(n) \le y_1}} \frac{\psi_p(n)}{n} + O(1) = \sum_{\substack{n \ge 1 \\ P^+(n) \le y_1}} \frac{\psi_p(n)}{n} + O(1) \\ &= \prod_{q \le y} \left(1 - \frac{1}{q}\right)^{-1} \exp\left(\sum_{y < q \le y_1} \frac{\psi_p(q)}{q} + O\left(\frac{1}{y}\right)\right) + O(1) \\ &\ge e^{\gamma} \log y \cdot \exp\left(- \sum_{y < q \le y_1} \frac{1}{q}\right)\left(1 + O\left(\frac{1}{\log y}\right)\right) + O(1) \\ &\ge e^{\gamma} \log y - e^{\gamma} \log\log y + O(1), \end{aligned}$$

by Mertens' theorem. Combining the above estimates and choosing $y = C\tau e^{\tau}$ for some large constant C, implies that $L(1, \psi_p) > e^{\gamma}\tau$ whenever $p \in \mathcal{P}(y)$. Appealing to (6-9) completes the proof. □

We end this section by proving the upper bound of (1-9) which we deduce from Theorem 2.7.

Proof of the upper bound of (1-9). Let $Z = x^{21/40}$ and $\delta = 1/100$. Let $y = e^{\tau - B}$, where B is a parameter to be chosen later. First by (2-9) we observe that

$$L(1, \psi_p) = \sum_{n \le Z} \frac{\psi_p(n)}{n} + O(x^{-\delta}) \le \sum_{\substack{n \ge 1 \\ P^+(n) \le y}} \frac{1}{n} + \Big| \sum_{\substack{1 \le n \le Z \\ P^+(n) > y}} \frac{\psi_p(n)}{n} \Big| + O(x^{-\delta})$$

$$\le e^{\gamma}\tau - e^{\gamma}B + \Big| \sum_{\substack{1 \le n \le Z \\ P^+(n) > y}} \frac{\psi_p(n)}{n} \Big| + O\Big(\frac{1}{\tau}\Big). \tag{6-10}$$

since $\prod_{p \le y}(1 - 1/p)^{-1} = e^{\gamma} \log y + O(1/\log y)$ by the prime number theorem. Therefore, there exists a positive constant c such that the proportion of primes $p \le x$ with $p \equiv a \mod 4$ and such that $L(1, \psi_p) \ge e^{\gamma}\tau$ is bounded by the proportion of primes $p \le x$ such that

$$\Big| \sum_{\substack{1 \le n \le Z \\ P^+(n) > y}} \frac{\psi_p(n)}{n} \Big| \ge e^{\gamma}\Big(B - \frac{c}{\tau}\Big).$$

Thus, appealing to Theorem 2.7 with $B = 2 + c/\tau$ completes the proof of (1-9) for $L(1, \psi_p)$. Finally, the analogous result for $L(1, \psi_p\chi_{-3})$ follows along the same lines upon taking $h(n) = \chi_{-3}(n)$ in Theorem 2.7 and noting that

$$\Big| \sum_{\substack{1 \le n \le Z \\ P^+(n) \le y}} \frac{\psi_p(n)\chi_{-3}(n)}{n} \Big| \le \prod_{\substack{q \ne 3 \\ q \le y}} \Big(1 - \frac{1}{q}\Big)^{-1} = \frac{2e^{\gamma}}{3} \log y + O\Big(\frac{1}{\log y}\Big),$$

by the prime number theorem. □

7. Positivity of sums of the Legendre symbol: Proof of Theorems 1.12 and 1.13

Let $Z = x^{21/40}$ and $\delta = 1/100$. Let $p \le x$ be a prime number such that $p \equiv 3 \mod 4$. Since ψ_p is odd and $\tau(\psi_p) = i\sqrt{p}$, then by (2-1) we have for any $\alpha \in (0, 1)$

$$\sum_{n \le \alpha p} \psi_p(n) = \frac{\sqrt{p}}{\pi} \sum_{1 \le n \le Z} \frac{\psi_p(n)(1 - \cos(2\pi n\alpha))}{n} + O\big(p^{1/2 - 1/40} \log p\big). \tag{7-1}$$

We shall now focus on the sum

$$F(\alpha, p) := \sum_{1 \le n \le Z} \frac{\psi_p(n)(1 - \cos(2\pi n\alpha))}{n}.$$

Our strategy for proving Theorems 1.12 and 1.13 consists in using Theorem 2.7 to bound the part of this sum over nonfriable integers uniformly over $\alpha \in [0, 1)$, for all primes $p \le x$ except for a small set of exceptions, and then prescribe the signs of $\psi_p(q)$ for the small primes q.

Proof of Theorem 1.12. We start by proving (1-10). Let $Z = x^{21/40}$ and $2 \le y \le \log x$ be a parameter to be chosen. Let $\mathcal{E}_1(x)$ be the set of primes $\sqrt{x} \le p \le x$ such that $p \equiv 3 \bmod 4$ and

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z\\ P^+(n)>y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} \right| > 2e^{\gamma}.$$

Then it follows from Theorem 2.7 that

$$|\mathcal{E}_1(x)| \ll \pi(x) \exp\left(-\frac{y}{3\log y}\right). \tag{7-2}$$

We now define $\mathcal{A}(y)$ to be the set of primes $\sqrt{x} \le p \le x$ such that $p \equiv 3 \bmod 4$ and $\psi_p(q) = 1$ for all primes $q \le y_0$, where $y_0 := y/(20\log y)$. Moreover, we put $\mathcal{D}(y) := \mathcal{A}(y) \setminus \mathcal{E}_1(x)$. By Lemma 6.2 and the estimate (7-2) we have $|\mathcal{D}(y)| \gg \pi(x)\exp(-y/(4\log y))$. Moreover, if $p \in \mathcal{D}(y)$ then it follows from (2-10) that for all $\alpha \in (0, 1)$ we have

$$\begin{aligned} F(\alpha, p) &= \sum_{\substack{1\le n\le Z\\ P^+(n)\le y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} + O(1) = \sum_{\substack{n\ge 1\\ P^+(n)\le y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} + O(1) \\ &= \sum_{\substack{n\ge 1\\ P^+(n)\le y_0}} \frac{1-\cos(2\pi n\alpha)}{n} + O(\log_2 y) = \sum_{\substack{n\ge 1\\ P^+(n)\le y}} \frac{1-\cos(2\pi n\alpha)}{n} + O(\log_2 y), \end{aligned} \tag{7-3}$$

since

$$\sum_{\substack{n\ge 1\\ y_0<P^+(n)\le y}} \frac{1}{n} = \prod_{p\le y}\left(1-\frac{1}{p}\right)^{-1} - \prod_{p\le y_0}\left(1-\frac{1}{p}\right)^{-1} \ll \log_2 y,$$

by Mertens' theorem. Furthermore, by Lemma 3.4 of [Bober et al. 2018] we have

$$\left| \sum_{\substack{n\ge 1\\ P^+(n)\le y}} \frac{\cos(2\pi n\alpha)}{n} \right| \le \left| \sum_{\substack{n\ge 1\\ P^+(n)\le y}} \frac{e(n\alpha)}{n} \right| = \sum_{\substack{n\le 1/\|\alpha\|\\ P^+(n)\le y}} \frac{1}{n} + O(1), \tag{7-4}$$

where $\|\cdot\|$ denotes the distance to the nearest integer. We now write $\|\alpha\| = y^{-u}$ for some $u > 0$. By Lemma 3.3 of [Bober et al. 2018] we get

$$\sum_{\substack{n\le 1/\|\alpha\|\\ P^+(n)\le y}} \frac{1}{n} = \sum_{\substack{n\le y^u\\ P^+(n)\le y}} \frac{1}{n} = (\log y)\int_0^u \rho(t)\,dt + O(1),$$

where ρ is the Dickman–de Bruijn function, defined by $\rho(t)$ for $0 \le t \le 1$, and $t\rho'(t) = -\rho(t-1)$ for $t > 1$. Combining this estimate with (7-3) and (7-4) we obtain

$$F(\alpha, p) \ge \log y \left(e^{\gamma} - \int_0^u \rho(t)\, dt \right) + O(\log_2 y), \tag{7-5}$$

for all $p \in \mathcal{D}(y)$ and all $\alpha \in (0,1)$, where $u = -\log\|\alpha\|/\log y$. Note that $\int_0^u \rho(t)\, dt$ is increasing in u and that $\int_0^\infty \rho(t)\, dt = e^{\gamma}$. Moreover, one has the estimate

$$e^{\gamma} - \int_0^u \rho(t)\, dt = \int_u^\infty \rho(t)\, dt = \frac{1}{u^{u(1+o(1))}}, \tag{7-6}$$

which follows from the standard estimate $\rho(t) = t^{-t(1+o(1))}$. Let C be a suitably large constant, and u_0 be the solution of the equation

$$\int_{u_0}^\infty \rho(t)\, dt = \frac{C \log_2 y}{\log y}. \tag{7-7}$$

Then, it follows from (7-5) that for all $p \in \mathcal{D}(y)$ and $\|\alpha\| > y^{-u_0}$ we have $F(\alpha, p) \ge 10$ if C is suitably large, and hence

$$\sum_{n \le \alpha p} \psi_p(n) > \sqrt{p}$$

by (7-1) if x is large enough. To finish the proof, we choose y such that $T = y^{u_0}/2$, which implies that $\lambda(p) > 1 - 1/T$ for all $p \in \mathcal{D}(y)$. Moreover, it follows from the estimates (7-6) and (7-7) that $u_0 = (1+o(1)) \log_2 y / \log_3 y$, and hence we get

$$y = \exp\left(\frac{(1+o(1)) \log T \log_3 T}{\log_2 T} \right),$$

as desired. □

In order to prove (1-11) for small T, we require a more precise estimate than the one provided by Lemma 6.2. To this end we prove the following result, which gives an asymptotic formula for the cardinality of the set $\mathcal{P}(x, y, a, \{\varepsilon_q\})$ in Lemma 6.2, in the smaller range $3 \le y \le \log\log x$.

Lemma 7.1. *Let $\{\varepsilon_q\}_{q \text{ prime}}$ be a sequence of ± 1, and $a \in \{1, 3\}$. Let x be large and $3 \le y \le \log\log x$ be a real number. Let $\mathcal{P}(x, y, a, \{\varepsilon_q\})$ be the set in Lemma 6.2. Then we have*

$$|\mathcal{P}(x, y, a, \{\varepsilon_q\})| = \frac{\pi(x)}{2^{\pi(y)+1}} + O(xe^{-c\sqrt{\log x}}),$$

for some positive constant c.

Proof. Let $Q = \prod_{q \le y} q$. Note that by the prime number theorem and our assumption we have $Q = e^{y+o(y)} \le (\log x)^{1+o(1)}$. We observe that for a prime $y < p \le x$ we have

$$\frac{1}{2^{\pi(y)+1}} (1 + \chi_{-4}(a)\chi_{-4}(p)) \prod_{q \le y} (1 + \varepsilon_q \psi_p(q)) = \begin{cases} 1 & \text{if } p \in \mathcal{P}(x, y, a, \{\varepsilon_q\}), \\ 0 & \text{otherwise,} \end{cases} \tag{7-8}$$

where χ_{-4} is the nonprincipal character modulo 4. We extend the sequence ε_q multiplicatively to all square-free numbers ℓ by letting $\varepsilon_\ell = \prod_{q|\ell} \varepsilon_q$. Therefore, expanding the product on the left-hand side of (7-8) we deduce that

$$\begin{aligned}|\mathcal{P}(x, y, a, \{\varepsilon_q\})| &= \frac{1}{2^{\pi(y)+1}} \sum_{y<p\le x} (1+\chi_{-4}(a)\chi_{-4}(p)) \prod_{q\le y} (1+\varepsilon_q \psi_p(q)) + O(y)\\ &= \frac{1}{2^{\pi(y)+1}} \sum_{p\le x} (1+\chi_{-4}(a)\chi_{-4}(p)) \sum_{\ell|Q} \varepsilon_\ell \psi_p(\ell) + O(y)\\ &= \frac{1}{2^{\pi(y)+1}} \left(\sum_{\ell|Q} \varepsilon_\ell \sum_{p\le x} \left(\frac{\ell}{p}\right) + \chi_{-4}(a) \sum_{\ell|Q} \varepsilon_\ell \sum_{p\le x} \chi_{-4}(p) \left(\frac{\ell}{p}\right)\right) + O(y). \end{aligned} \tag{7-9}$$

If $\ell \neq 1$, then the law of quadratic reciprocity implies that $\xi_1 = \left(\frac{\ell}{\cdot}\right)$ is a nonprincipal character of conductor ℓ or 4ℓ. Similarly, for all ℓ the character $\xi_2 = \chi_{-4}\xi_1$ is a nonprincipal character of conductor 4ℓ. Furthermore, note that $4\ell \le 4Q \le (\log x)^2$ if x is large enough. Thus, it follows from Corollary 11.18 of [Montgomery and Vaughan 2007] that for $j = 1$ if $\ell \neq 1$ and $\ell \mid Q$, and for $j = 2$ for all $\ell \mid Q$ we have

$$\sum_{p\le x} \xi_j(p) \ll x \exp(-c\sqrt{\log x}). \tag{7-10}$$

Inserting these bounds in (7-9) completes the proof. □

Proof of Theorem 1.13. We start by proving (1-11). Let $Z = x^{21/40}$ and $\delta = 1/100$. By (7-1) we have

$$\sum_{n\le \alpha p} \psi_p(n) = \frac{\sqrt{p}}{\pi} F(\alpha, p) + O(p^{1/2-\delta}),$$

for all primes $\sqrt{x} \le p \le x$ and all $\alpha \in (0, 1)$. Let $3 \le y \le (\log_2 x)(\log_3 x)^2$ be a parameter to be chosen, and $\mathcal{E}_2(x)$ be the set of primes $\sqrt{x} \le p \le x$ such that $p \equiv 3 \bmod 4$ and

$$\max_{\alpha\in[0,1)} \left| \sum_{\substack{1\le n\le Z\\ P^+(n)>y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} \right| > \frac{2e^\gamma}{\log y}.$$

By Theorem 2.7 we have

$$|\mathcal{E}_2(x)| \ll \pi(x) \exp\left(-\frac{y}{3(\log y)^3}\right). \tag{7-11}$$

We will use the same choice as Montgomery [1976] for the values of $\psi_p(q)$ for small q. Let $h(n)$ be the completely multiplicative function such that $h(q) = \chi_{-3}(q)$ for all primes $q \neq 3$ and $h(3) = -1$. Montgomery [1976] showed that for all $\alpha \in (0, 1)$ such that $\|\alpha\| < 1/3$ we have

$$U(\alpha) := \sum_{n=1}^{\infty} \frac{h(n)\cos(2\pi n\alpha)}{n} > \frac{\pi}{8\sqrt{3}}. \tag{7-12}$$

Let $y_0 := y/(3\log y)^2$ and $3 \le H \le y_0$ be a parameter to be chosen. Let $\mathcal{B}(y, H)$ be the set of primes $\sqrt{x} \le p \le x$ such that $p \equiv 3 \bmod 4$ and

$$\psi_p(q) = \begin{cases} h(q) & \text{if } q \le H, \\ -1 & \text{if } H < q \le y_0. \end{cases}$$

We define $\mathcal{T}(y) := \mathcal{B}(y, H) \setminus \mathcal{E}_2(x)$. Then we have $|\mathcal{T}(y)| \gg \pi(x)/2^{\pi(y_0)}$ by Lemma 7.1 and the estimate (7-11). We now let p be a prime in $\mathcal{T}(y)$. By (2-10) and our assumption on p we get

$$\begin{aligned} F(\alpha, p) &= \sum_{\substack{1 \le n \le Z \\ P^+(n) \le y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} + O\Big(\frac{1}{\log y}\Big) \\ &= \sum_{\substack{n \ge 1 \\ P^+(n) \le y}} \frac{\psi_p(n)(1-\cos(2\pi n\alpha))}{n} + O\Big(\frac{1}{\log y}\Big). \end{aligned} \tag{7-13}$$

Furthermore, we have

$$\begin{aligned} \sum_{\substack{n \ge 1 \\ P^+(n) \le y}} \frac{\psi_p(n)}{n} &= \prod_{q \le y} \Big(1 - \frac{\psi_p(q)}{q}\Big)^{-1} \\ &= \frac{3}{4} \prod_{q \le H} \Big(1 - \frac{\chi_{-3}(q)}{q}\Big)^{-1} \prod_{H < q \le y_0} \Big(1 + \frac{1}{q}\Big)^{-1} \prod_{y_0 < q \le y} \Big(1 - \frac{\psi_p(q)}{q}\Big)^{-1}. \end{aligned} \tag{7-14}$$

By the prime number theorem in arithmetic progressions we have

$$\sum_{q \le t} \chi_{-3}(q) \ll t \exp(-c\sqrt{\log t}),$$

for some constant $c > 0$. Therefore, by partial summation we obtain

$$\sum_{q > H} \frac{\chi_{-3}(q)}{q} \ll \exp\Big(-\frac{c}{2}\sqrt{\log H}\Big).$$

This implies that

$$\begin{aligned} \prod_{q \le H} \Big(1 - \frac{\chi_{-3}(q)}{q}\Big)^{-1} &= L(1, \chi_{-3}) + O\Big(\exp\Big(-\frac{c}{2}\sqrt{\log H}\Big)\Big) \\ &= \frac{\pi}{3\sqrt{3}} + O\Big(\exp\Big(-\frac{c}{2}\sqrt{\log H}\Big)\Big). \end{aligned}$$

Inserting this estimate in (7-14) and using Mertens' theorem we deduce that

$$0 < \sum_{\substack{n \ge 1 \\ P^+(n) \le y}} \frac{\psi_p(n)}{n} \le \frac{\pi}{4\sqrt{3}} \frac{\log H \log y}{(\log y_0)^2} \Big(1 + O\Big(\frac{1}{\log H}\Big)\Big).$$

We now choose $H = \sqrt{y}/(C_1(\log y)^2)$ where C_1 is a suitably large constant. This gives

$$\sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)}{n} \leq \frac{\pi}{8\sqrt{3}} - \frac{\log C_1}{\log y}. \tag{7-15}$$

On the other hand, by Parseval's identity we have

$$\begin{aligned}\int_0^1 \Bigg| \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)\cos(2\pi n\alpha)}{n} - U(\alpha)\Bigg|^2 d\alpha &= \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{(\psi_p(n)-h(n))^2}{n^2} + \sum_{\substack{n\geq 1\\ P^+(n)> y}} \frac{h(n)^2}{n^2}\\ &\ll \sum_{\substack{n\geq 1\\ P^+(n)> H}} \frac{1}{n^2} \ll \frac{\log y}{\sqrt{y}},\end{aligned}$$

since

$$\sum_{\substack{n\geq 1\\ P^+(n)> H}} \frac{1}{n^2} = \frac{\pi^2}{6} - \sum_{\substack{n\geq 1\\ P^+(n)\leq H}} \frac{1}{n^2} = \frac{\pi^2}{6}\Bigg(1 - \prod_{q>H}\Big(1-\frac{1}{q^2}\Big)\Bigg) \ll \sum_{q>H}\frac{1}{q^2} \ll \frac{1}{H\log H}.$$

Let $\mathcal{S}_p$ be the set of $\alpha \in (0,1)$ such that

$$\Bigg| \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)\cos(2\pi n\alpha)}{n} - U(\alpha)\Bigg| > \frac{1}{\log y}.$$

Then

$$\mu(\mathcal{S}_p) \leq (\log y)^2 \int_0^1 \Bigg| \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)\cos(2\pi n\alpha)}{n} - U(\alpha)\Bigg|^2 d\alpha \ll \frac{(\log y)^3}{\sqrt{y}}.$$

Thus, recalling the definitions of the sets $\mathcal{T}(y)$ and $\mathcal{S}_p$ and using the estimate (7-12) we deduce that if $p \in \mathcal{T}(y)$ and $\alpha \in (0,1/3)\cup(2/3,1)\setminus \mathcal{S}_p$ then

$$\sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)\cos(2\pi n\alpha)}{n} > \frac{\pi}{8\sqrt{3}} - \frac{1}{\log y}.$$

Combining this estimate with (7-13) and (7-15) gives

$$F(\alpha,p) = \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)}{n} - \sum_{\substack{n\geq 1\\ P^+(n)\leq y}} \frac{\psi_p(n)\cos(2\pi n\alpha)}{n} + O\Big(\frac{1}{\log y}\Big) \leq -\frac{C_2}{\log y},$$

for some positive constant C_2 if C_1 is suitably large. This in turn implies that

$$\sum_{n\leq \alpha p} \psi_p(n) < -\frac{\sqrt{p}}{\log_3 p}$$

by (7-1), if x is large enough. Hence for $p \in \mathcal{T}(y)$ we have

$$\lambda(p) \le \frac{1}{3} + \mu(\mathcal{S}_p) \le \frac{1}{3} + C_3 \frac{(\log y)^3}{\sqrt{y}},$$

for some absolute constant $C_3 > 0$. Choosing $y = C_4 T^2 (\log T)^6$ for some suitably large constant C_4 completes the proof of (1-11).

To prove (1-12) we follow the exact same lines, and replace Lemma 7.1 by Lemma 6.2. In this case we make the following choices of the parameters: $y_0 = y/(3\log y)^3 \le c_0 \log x$ (where c_0 is the constant in Lemma 6.2), $H = \sqrt{y}/(C_5(\log y)^3)$, and $y = C_6 T^2 (\log T)^8$ for some suitably large constants C_5 and C_6. This completes the proof. □

Acknowledgments

The author would like to thank Andrew Granville for useful discussions. We would also like to thank the referee for carefully reading the paper and for comments and suggestions. This work was completed while the author was on a Délégation CNRS at the IRL3457 CRM-CNRS in Montréal. We would like to thank the CNRS for its support and the Centre de Recherches Mathématiques for its excellent working conditions.

References

[Baker and Montgomery 1990] R. C. Baker and H. L. Montgomery, "Oscillations of quadratic L-functions", pp. 23–40 in *Analytic number theory* (Allerton Park, IL, 1989), edited by B. C. Berndt et al., Progr. Math. **85**, Birkhäuser, Boston, MA, 1990. MR Zbl

[Bateman and Chowla 1950] P. T. Bateman and S. Chowla, "Averages of character sums", *Proc. Amer. Math. Soc.* **1** (1950), 781–787. MR Zbl

[Bober 2014] J. Bober, "Averages of character sums", preprint, 2014. arXiv 1409.1840

[Bober and Goldmakher 2013] J. W. Bober and L. Goldmakher, "The distribution of the maximum of character sums", *Mathematika* **59**:2 (2013), 427–442. MR Zbl

[Bober and Goldmakher 2016] J. W. Bober and L. Goldmakher, "Pólya–Vinogradov and the least quadratic nonresidue", *Math. Ann.* **366**:1-2 (2016), 853–863. MR Zbl

[Bober et al. 2018] J. Bober, L. Goldmakher, A. Granville, and D. Koukoulopoulos, "The frequency and the structure of large character sums", *J. Eur. Math. Soc.* **20**:7 (2018), 1759–1818. MR Zbl

[Bombieri 1987] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, Astérisque **18**, Soc. Math. France, Paris, 1987. MR Zbl

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Grad. Texts in Math. **74**, Springer, 2000. MR Zbl

[Elliott 1970] P. D. T. A. Elliott, "On the mean value of $f(p)$", *Proc. Lond. Math. Soc.* (3) **21**:1 (1970), 28–96. MR Zbl

[Goldmakher 2012] L. Goldmakher, "Multiplicative mimicry and improvements to the Pólya–Vinogradov inequality", *Algebra Number Theory* **6**:1 (2012), 123–163. MR Zbl

[Granville and Mangerel 2023] A. Granville and A. P. Mangerel, "Three conjectures about character sums", *Math. Z.* **305**:3 (2023), art. id. 49. MR Zbl

[Granville and Soundararajan 2003] A. Granville and K. Soundararajan, "The distribution of values of $L(1, \chi_d)$", *Geom. Funct. Anal.* **13**:5 (2003), 992–1028. MR Zbl

[Granville and Soundararajan 2006] A. Granville and K. Soundararajan, "Extreme values of $|\zeta(1+it)|$", pp. 65–80 in *The Riemann zeta function and related themes* (Bangalore, 2003), edited by R. Balasubramanian and K. Srinivas, Ramanujan Math. Soc. Lect. Notes Ser. **2**, Ramanujan Math. Soc., Mysuru, India, 2006. MR Zbl

[Granville and Soundararajan 2007] A. Granville and K. Soundararajan, "Large character sums: pretentious characters and the Pólya–Vinogradov theorem", *J. Amer. Math. Soc.* **20**:2 (2007), 357–384. MR Zbl

[Heath-Brown 1995] D. R. Heath-Brown, "A mean value estimate for real character sums", *Acta Arith.* **72**:3 (1995), 235–275. MR Zbl

[Holmin et al. 2019] S. Holmin, N. Jones, P. Kurlberg, C. McLeman, and K. Petersen, "Missing class groups and class number statistics for imaginary quadratic fields", *Exp. Math.* **28**:2 (2019), 233–254. MR Zbl

[Hussain 2022a] A. Hussain, *Character studies: investigating the limiting distribution of character sums*, Ph.D. thesis, University of Bristol, 2022, available at https://research-information.bris.ac.uk/en/studentTheses/character-studies.

[Hussain 2022b] A. Hussain, "The limiting distribution of character sums", *Int. Math. Res. Not.* **2022**:20 (2022), 16292–16326. MR Zbl

[Hussain and Lamzouri 2023] A. Hussain and Y. Lamzouri, "The limiting distribution of Legendre paths", 2023. To appear in *J. Éc. Polytech. Math.* arXiv 2304.13025

[Joshi 1970] P. T. Joshi, "The size of $L(1, \chi)$ for real nonprincipal residue characters χ with prime modulus", *J. Number Theory* **2**:1 (1970), 58–73. MR Zbl

[Lamzouri 2010] Y. Lamzouri, "Distribution of values of L-functions at the edge of the critical strip", *Proc. Lond. Math. Soc.* (3) **100**:3 (2010), 835–863. MR Zbl

[Lamzouri 2011] Y. Lamzouri, "Extreme values of $\arg L(1, \chi)$", *Acta Arith.* **146**:4 (2011), 335–354. MR Zbl

[Lamzouri 2017] Y. Lamzouri, "The number of imaginary quadratic fields with prime discriminant and class number up to H", *Q. J. Math.* **68**:4 (2017), 1379–1393. MR Zbl

[Lamzouri and Mangerel 2022] Y. Lamzouri and A. P. Mangerel, "Large odd order character sums and improvements of the Pólya–Vinogradov inequality", *Trans. Amer. Math. Soc.* **375**:6 (2022), 3759–3793. MR Zbl

[Littlewood 1928] J. E. Littlewood, "On the class number of the corpus $P(\sqrt{-k})$", *Proc. Lond. Math. Soc.* (2) **27**:5 (1928), 358–372. MR Zbl

[Mangerel 2020] A. P. Mangerel, "Short character sums and the Pólya–Vinogradov inequality", *Q. J. Math.* **71**:4 (2020), 1281–1308. MR Zbl

[Mehkari 2005] S. Mehkari, *Distribution of sums of the Legendre symbol*, master's thesis, Université de Montréal, 2005, available at https://hdl.handle.net/1866/17278.

[Montgomery 1971] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math. **227**, Springer, 1971. MR Zbl

[Montgomery 1976] H. L. Montgomery, "Distribution questions concerning a character sum", pp. 195–203 in *Topics in number theory* (Debrecen, Hungary, 1974), edited by P. Turán, Colloq. Math. Soc. János Bolyai **13**, North-Holland, Amsterdam, 1976. MR Zbl

[Montgomery and Vaughan 1977] H. L. Montgomery and R. C. Vaughan, "Exponential sums with multiplicative coefficients", *Invent. Math.* **43**:1 (1977), 69–82. MR Zbl

[Montgomery and Vaughan 1979] H. L. Montgomery and R. C. Vaughan, "Mean values of character sums", *Canadian J. Math.* **31**:3 (1979), 476–487. MR Zbl

[Montgomery and Vaughan 1999] H. L. Montgomery and R. C. Vaughan, "Extreme values of Dirichlet L-functions at 1", pp. 1039–1052 in *Number theory in progress*, *II* (Zakopane/Kościelisko, Poland, 1997), edited by K. Győry et al., de Gruyter, Berlin, 1999. MR Zbl

[Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Stud. Adv. Math. **97**, Cambridge Univ. Press, 2007. MR Zbl

[Paley 1932] R. E. A. C. Paley, "A theorem on characters", *J. Lond. Math. Soc.* **7**:1 (1932), 28–32. MR Zbl

Communicated by Roger Heath-Brown
Received 2023-03-02 Revised 2023-09-12 Accepted 2023-11-27

youness.lamzouri@univ-lorraine.fr *Institut Élie Cartan de Lorraine, Université de Lorraine, CNRS, Nancy, France*

IRL3457 CRM-CNRS, Centre de Recherches Mathématiques, Université de Montréal, Montréal, QC, Canada

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 18 No. 11 2024