THE LINEAR INDEPENDENCE OF 1, $\zeta(2)$, AND $L(2,\chi_{-3})$

FRANK CALEGARI, VESSELIN DIMITROV, AND YUNQING TANG

Abstract. We prove the irrationality of the classical Dirichlet L-value

$$L(2,\chi_{-3}) = \frac{1}{1^2} - \frac{1}{2^2} + \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{7^2} - \frac{1}{8^2} + \ldots.$$

The argument applies a new kind of arithmetic holonomy bound to a well-known construction of Zagier [Zag09]. In fact our work also establishes the $\mathbf{Q}$-linear independence of $1, \zeta(2)$ and $L(2,\chi_{-3})$. We also give a number of other applications of our method to other problems in irrationality.

Contents

1. Introduction

1.1. **Dirichlet L-values.**

The values of the Riemann zeta function $\zeta(k)$ for positive integers k, and more generally the Dirichlet L-values

$$L(k,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^k}$$

for quadratic characters χ, have long been a source of interest to mathematicians. Suppose that χ is a primitive quadratic character of conductor D, where we use the convention that the sign of D is the sign of $\chi(-1)$. Starting with work of Euler [Eul1735] and Dirichlet [Dir1837] (or even far before that in the special case of $D=-4$ and $k=1$ [Roy90]), we know that, for positive integers k:

$$L(k,\chi) \in \begin{cases} \pi^k \cdot \sqrt{D} \cdot \mathbf{Q}^\times, & k \text{ even and } \chi(-1)=1, \\ \pi^k \cdot \sqrt{-D} \cdot \mathbf{Q}^\times, & k \text{ odd and } \chi(-1)=-1, \\ \sqrt{D} \log\left(|\overline{\mathbf{Q}}^\times| \smallsetminus \{1\}\right), & k=1,\ \chi(-1)=1, \text{ and } D \neq 1. \end{cases} \tag{1.1.1}$$

Combined with Lindemann's theorem [Lin1882] that π is transcendental and Weierstrass's extension [Wei1885] to the transcendence of the natural logarithms of algebraic numbers other than 0 or 1, one knows all of these values to be transcendental. The remaining L-values are far less well understood. Indeed, in the (approximately) last 140 years since [Wei1885] only a *single* further explicit number $L(k,\chi)$ has been shown to be irrational, namely Apéry's unexpected 1978 proof that $\zeta(3)$ is irrational [Ape79, vdP79, Coh78]. In this paper, we establish the irrationality of a new L-value $L(k,\chi)$; in some sense the "simplest" open case corresponding to $k=2$ and the character $\chi=\chi_{-3}$ of smallest possible conductor:

Theorem A. *The period*

$$\begin{aligned} L(2,\chi_{-3}) &= \sum_{n=0}^{\infty} \left(\frac{1}{(3n+1)^2} - \frac{1}{(3n+2)^2} \right) = 0.7813024128964862968\ldots \\ &= \iint_{1 \geq y \geq x \geq 0} \frac{dx\,dy}{y(1+x+x^2)} = -\int_0^1 \frac{\log(x)dx}{1+x+x^2} \end{aligned}$$

is irrational. More generally, the three periods $1, \pi^2, L(2,\chi_{-3})$ are linearly independent over $\mathbf{Q}$.

The formula above exhibits $L(2,\chi_{-3})$ as a period in the sense of Kontsevich–Zagier [KZ01]. There are a panoply of other more complicated expressions for $L(2,\chi_{-3})$ as an integral, or an infinite sum, for example, the following sum of hypergeometric type ([HPHP11, §3]):

$$L(2,\chi_{-3}) = \frac{1}{27} \sum_{n=1}^{\infty} \frac{(4-15n)(-27)^n}{n^3 \binom{2n}{n}^2 \binom{3n}{n}},$$

or, more serendipitously, in terms of the sum of the inverse squares of the entries greater than one in Pascal's triangle [Sta23]:

$$L(2,\chi_{-3}) = -\frac{1}{3} + \sum_{n>m>0} \frac{1}{\binom{n}{m}^2}.$$

The constant $3\sqrt{3}\,L(2,\chi_{-3})/4 = \mathrm{Im}(\mathrm{Li}_2(e^{\pi i/3})) = \frac{3}{2}\cdot\mathrm{Im}(\mathrm{Li}_2(e^{2\pi i/3})) = 1.014941\ldots$ is the volume of the regular ideal hyperbolic tetrahedron (the one of the maximal volume), and is also the volume of the non-compact hyperbolic manifold with minimal volume [Ada87] (the Gieseking manifold, whose orientable double cover is the complement of the figure 8 knot [Thu97]). It is an open problem to show that the volumes of hyperbolic 3-manifolds are not all rationally related (see [Thu82, Problem 23], and [Mil82, Mil83]). While our result does not have any direct implications for this question, it is the first unconditional result to make contact with the arithmetic nature of these volumes. Another appearance of $L(2,\chi_{-3})$ is in Smyth's formula [BZ20, Prop. 3.4]

$$\frac{3\sqrt{3}}{4\pi}L(2,\chi_{-3}) = m(1+x+y) := \int_0^1\int_0^1 \log|1+e^{2\pi is}+e^{2\pi it}|\,ds\,dt \qquad (1.1.2)$$

linking $L(2,\chi_{-3})$ to the *Mahler measure* of the simplest essentially bivariate polynomial $1+x+y$, or equivalently in the language of Diophantine and Arakelov geometry [Phi91, BGS94], to the canonical height of the subvariety $1+x+y=0$ of the linear algebraic torus $\mathbf{G}_m^2$. Unfortunately, while the nonvanishing Mahler measures of the integer univariate polynomials are all known to be transcendental by the Hermite–Lindemann–Weierstrass theorem [Her1874, Lin1882, Wei1885], our result has no direct bearing on the conjectured irrationality of any such canonical heights! (We do, incidentally, also prove the irrationality of the Mahler measure of the *rational* coefficients bivariate polynomial $(1+x+y)^4/3$, which is not a canonical height. This is in Theorem 2.11.17.)

An immediate consequence of Theorem A is the irrationality of the following values of the "trigamma" function $\psi_1(z) = \sum_{n=0}^{\infty}\frac{1}{(z+n)^2} = \zeta(2,z)$:

Corollary B. *The following numbers are irrational:*

$$\begin{aligned}
\frac{1}{1^2}+\frac{1}{4^2}+\frac{1}{7^2}+\frac{1}{10^2}+\ldots &= \frac{L(2,\chi_{-3})}{2}+\frac{2\pi^2}{27},\\
\frac{1}{2^2}+\frac{1}{5^2}+\frac{1}{8^2}+\frac{1}{11^2}+\ldots &= -\frac{L(2,\chi_{-3})}{2}+\frac{2\pi^2}{27},\\
\frac{1}{1^2}+\frac{1}{7^2}+\frac{1}{13^2}+\frac{1}{19^2}+\ldots &= \frac{5L(2,\chi_{-3})}{8}+\frac{\pi^2}{18},\\
\frac{1}{5^2}+\frac{1}{11^2}+\frac{1}{17^2}+\frac{1}{23^2}+\ldots &= -\frac{5L(2,\chi_{-3})}{8}+\frac{\pi^2}{18}.
\end{aligned}$$

Note that since $\psi_1(x+1)-\psi_1(x)+1/x^2=0$, it also follows from Theorem A (together with the fact that $\psi_1(1)=\pi^2/6$ and $\psi_1(1/2)=\pi^2/2$) that $\psi_1(n/6)$ is irrational for any $n\in\mathbf{N}_{>0}$. These are the first new irrationality results for ψ_1 since Legendre's proof in 1794 [Leg1794] that π^2 is irrational!

As another application of what turns out to be exactly the same argument, we also prove the following irrationality result for certain products of two logarithms (see Theorem 14.0.1 in § 14).

Theorem C. *Let $m,n\in\mathbf{Z}\smallsetminus\{-1,0\}$ be integers such that $\left|\frac{m}{n}-1\right|<\frac{1}{10^6}$. Then*

$$\log\left(1+\frac{1}{m}\right)\log\left(1+\frac{1}{n}\right)$$

is irrational. Moreover, for $m \neq n$, the following are linearly independent over $\mathbf{Q}$*:*

$$1, \quad \log\left(1+\frac{1}{m}\right), \quad \log\left(1+\frac{1}{n}\right), \quad \log\left(1+\frac{1}{m}\right)\log\left(1+\frac{1}{n}\right).$$

1.2. **Comparisons to the work of Apéry.** Apéry's proof [vdP79] consisted of finding an explicit sequence of rational approximations which converged "sufficiently quickly" to $\zeta(3)$ to prove that $\zeta(3)$ is irrational. Ever since Apéry's result, considerable effort has been expended in searching for analogous sequences which demonstrate the irrationality of other L-values $L(k,\chi)$ beyond those of the form (1.1.1). Unfortunately, despite enormous efforts, no such sequences have ever been found.[1] In particular, in this paper, we do *not* find (directly) any new convergent sequences to $L(2,\chi_{-3})$. Instead, we show how one can exploit the arithmetic nature of *known* approximations (found by Apéry and others) in a more subtle way using both methods from transcendental number theory and complex analysis.

In order to introduce our main idea, we begin with an exposition and then a reformulation of some of the key features of Apéry's proof. The first remark to make is that Apéry's proof uses very little number theory; indeed the only number theoretic input is a (weak form) of the prime number theorem and the following elementary lemma:

Lemma 1.2.1. *If there is a $\delta > 0$ and a sequence of rational numbers $p_n/q_n \neq \beta$ with $q_n \to \infty$ such that*

$$\left|\beta - \frac{p_n}{q_n}\right| < \frac{1}{q_n^{1+\delta}} \quad n = 1, 2, \ldots,$$

then β is irrational.

This lemma is true even with the weaker hypothesis that $|\beta - p_n/q_n| = o(1/q_n)$. Apéry writes down a pair of power series $A(x), B(x) \in \mathbf{Q}[\![x]\!]$ and the linear combination

$$P(x) = B(x) - \zeta(3)A(x) = \sum_{n=0}^{\infty} x^n \left(b_n - \zeta(3)a_n\right).$$

The coefficients a_n and b_n are rational numbers, and more precisely:

$$a_n \in \mathbf{Z}, \qquad [1,2,3,\ldots,n]^3 b_n \in \mathbf{Z}.$$

Here and throughout our paper, we follow the conventional notation $[1,2,\ldots,n]$ for the lowest common multiple of the first n integers. The prime number theorem determines the growth rate of these denominators:

$$\log[1,2,3,\ldots,n] = n + o(n).$$

At the same time, Apéry proves that $A(x)$ (and $B(x)$) have radius of convergence $(\sqrt{2}-1)^4$ whereas $P(x)$ has radius of convergence exactly $(\sqrt{2}+1)^4$. Now one exploits the inequality

$$4\log(\sqrt{2}+1) > 3 \tag{1.2.2}$$

[1]One significant step which is not directly related to the irrationality of specific L-values is the theorem of Ball and Rivoal [Riv00, BR01] that infinitely many odd zeta values are irrational. One refinement by Zudilin [Zud01] proves that at least one of the values $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ and $\zeta(11)$ is irrational.

to deduce, by Lemma 1.2.1 with $p_n/q_n = b_n/a_n$ and

$$\delta = \frac{4\log(\sqrt{2}+1)-3}{4\log(\sqrt{2}+1)+3},$$

that $\zeta(3) \notin \mathbf{Q}$.

There are a number of other situations where one can construct functions $A(x)$, $B(x)$ of a similar flavour so that a particular linear combination $P(x) = B(x) - \eta A(x)$ has extra convergence properties, and where $\eta = L(k,\chi)$ turns out to be the unique complex number characterized in this way. But the analogous inequality (1.2.2) always seems to fail,[2] and one can draw no consequences about the arithmetic of the corresponding L-value. (For one particularly interesting study of sequences of the form considered by Apéry, see [Zag09]. In our proof of Theorem A, we will make a central use of some of the sequences (re-)discovered by Zagier in his search.)

The starting point of our investigation is that, even when the analogue of (1.2.2) fails as it usually does, the functions $P(x)$ arising in these constructions have more structure which has not previously been exploited. Apéry's functions $A(x)$ and $B(x)$ turn out to satisfy a linear ordinary differential equation (ODE) with coefficients in $\mathbf{Z}[x]$ which only has (regular) singular points at $x = 0$, ∞, and $(\sqrt{2}\pm 1)^4$. The function $P(x)$ arises as the unique (up to scalar) linear combination of the two dimensional space of solutions to this ODE which are holomorphic at 0 with the additional property that it is also holomorphic at $(\sqrt{2}-1)^4$. This implies that $P(x)$, for example, is not merely holomorphic on the disc of radius $(\sqrt{2}+1)^4$, but extends to a holomorphic function on all of $\mathbf{C} \smallsetminus [(\sqrt{2}+1)^4, \infty)$, or (more relevantly for our ultimate purposes, but less important for the introduction) to a function on the universal cover of

$$\mathbf{P}^1 \smallsetminus \{0, (\sqrt{2}-1)^4, (\sqrt{2}+1)^4, \infty\}$$

which is holomorphic at $x = 0$ and overconverges beyond the first singularity $(\sqrt{2}-1)^4$. All Apéry uses is that $P(x)$ is holomorphic on the disc of radius $(\sqrt{2}+1)^4$.

Now imagine an analogous situation where $A(x)$ and $B(x)$ are holomorphic (at $x = 0$) solutions to an ODE[3] with regular singular points at 0, ∞ and a pair of real numbers $0 < \alpha < \beta$, and $P(x) = A(x) - \eta B(x)$ is a linear combination which is also holomorphic at α, but whose analytic continuation has a singularity at β and does not analytically continue to a meromorphic function at $x = \beta$. But now suppose — taking into account the denominators of a_n and b_n — that the constant β is not large enough to imply that the corresponding convergents p_n/q_n satisfy Lemma 1.2.1. Is there a way to exploit the fact that not only is $P(x)$ a holomorphic function on the disc of radius β, but also extends to a holomorphic function on $\mathbf{C} \smallsetminus [\beta, \infty)$ and on the universal cover of $\mathbf{P}^1 \smallsetminus \{0, \alpha, \beta, \infty\}$?

To make things simpler (too simple, in fact — we will return to the issue of the necessity of denominators), let us momentarily suppose that the a_n and b_n are actually integers. To run Apéry's proof scheme via Lemma 1.2.1, it then would have sufficed that $\beta > 1$. Lemma 1.2.1 in this case has the following alternate formulation:

[2]except in one notable example found by Apéry himself with $\eta = \zeta(2)$.

[3]By this we mean: a *linear* ODE over $\mathbf{Q}(x)$, as always in this paper.

Lemma 1.2.3 (Simple Lemma)**.** *An integer power series in $\mathbf{Z}[\![x]\!]$ that defines a holomorphic function on the disc $|x| < R$ of a radius $R > 1$ is a polynomial.*

In our running example, if $\beta < 1$, we can deduce nothing from Lemma 1.2.3, but we can still derive that $P(x)$ is holomorphic on $\mathbf{C} \smallsetminus [\beta, \infty)$. There is an entire subject devoted to more subtle extensions of Lemma 1.2.3, beginning with the Theorem of Borel–Pólya [Ami75, Chapter 5], which allows one to make conclusions about $P(x) \in \mathbf{Z}[\![x]\!]$ from weaker analytic hypotheses than simply converging on a disc of sufficiently large radius. We recall (a special case of) this theorem now. If $\Omega \subset \mathbf{C}$ is a simply connected open region containing 0, then, from the Riemann mapping theorem, there exists a biholomorphic map $\varphi : \mathbf{D} \to \Omega$ with $\varphi(0) = 0$. The map φ is unique up to biholomorphisms of the unit disc fixing 0, which are all given by rotations. In particular, the invariant $|\varphi'(0)|$ does not depend on the choice of φ, and (by definition) is equal to the conformal radius $\rho(\Omega, 0)$ of Ω at 0. The conformal radius of the disc $\mathbf{D}_R = D(0, R)$ is equal to R (via the map $\varphi(z) = Rz$), but the conformal radius of any other Ω is strictly larger than the radius of the largest disc contained in Ω and centered at $z = 0$. We have:

Theorem 1.2.4 (Borel–Pólya, [Pól1923])**.** *A power series $P(x) \in \mathbf{Z}[\![x]\!]$ that continues analytically to a simply connected open region $0 \in \Omega \subset \mathbf{C}$ of conformal radius $\rho(\Omega, 0) = |\varphi'(0)| > 1$ is necessarily a rational function: $P(x) \in \mathbf{Q}(x)$.*

For example, the biholomorphic map

$$\varphi : \mathbf{D} \to \mathbf{C} \smallsetminus [\beta, \infty), \quad z \mapsto \frac{4\beta z}{(1+z)^2}$$

shows that $\mathbf{C} \smallsetminus [\beta, \infty)$ has conformal radius $|\varphi'(0)| = 4\beta$. It follows from Theorem 1.2.4 in our imagined example above that $P(x)$ is a rational function as soon as $4\beta > 1$, contradicting the assumption that $P(x)$ was not meromorphic at β, and implying that η is irrational. This is already clearly an improvement on the condition that $\beta > 1$. (The basic idea for this special case of Pólya's theorem is sketched in Remark 1.2.5 at the end of this introduction.)

Even beyond Theorem 1.2.4 (as we shall discuss in Section 2 below), there are algebraicity theorems of André and others with even weaker hypotheses that allow one to deduce that $P(x)$ is algebraic over $\mathbf{Q}(x)$ (see [And04, And89] and [CDT21]), which can often be ruled out directly in practice for any particular $P(x)$.

The main thrust of our paper is in adapting and honing up the methods of Borel, Pólya, and André to fit into the Apéry irrationality proofs context. The algebraicity criteria as such do not apply, because the power series $A(x)$ and $B(x)$ of relevance to Apéry style proofs never (both) have integral coefficients. And indeed, when one introduces denominators (even of some controlled flavour), it turns out that algebraicity is no longer the right property to consider. To begin with, a theorem of Eisenstein [BG06, §11.4] states that the power series expansion of any algebraic function in $\overline{\mathbf{Q}(x)} \cap \mathbf{Q}[\![x]\!]$ has $\mathbf{Z}[1/S]$ coefficients for some $S \in \mathbf{N}_{>0}$. But from the point of view of the various proofs of Borel's theorem (and its variations), if $P(x) \in \mathbf{Z}[\![x]\!]$, then so too are all of its powers; but if $P(x)$ has denominators, then the powers of $P(x)$ typically have *worse* denominators. An example to keep in mind is

$$-\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}.$$

This function has the property that multiplication by $[1, 2, 3, \ldots, n]$ (of order e^n) simultaneously clears the denominators of the first n coefficients. But in order to clear the denominators of the first n coefficients of $\log^m(1-x)$, one has to multiply by a denominator of order

$$[1, 2, \ldots, n] \times [1, \ldots, n/2] \times \cdots \times [1, \ldots, n/m] = \exp\left(n\left(1 + \frac{1}{2} + \ldots + \frac{1}{m}\right) + o(n)\right).$$

Here and throughout the paper by $[1, 2, 3, \ldots, bn]$ for $b \in \mathbf{R}^{>0}$ we mean by abuse of notation $[1, 2, 3, \ldots, \lfloor bn \rfloor]$. On the other hand, *differentiation* does preserve the property of controlled denominator growth. Hence, instead of an algebraicity theorem, one should expect an *arithmetic holonomy* bound, where one bounds the dimension of a $\mathbf{Q}(x)$-vector space generated by functions with certain denominator growth and analytic properties, and which is closed under differentiation. This in particular implies that, in the appropriate generalization of the Borel–Pólya conditions, the *solutions* — a precise formulation is given by Corollary 2.6.1, to be discussed in detail in the next section — are G-functions in the sense of Siegel (see [Zan14] and [DGS94, §VIII.1], see also Definition 15.1.1). Moreover, one can hope to give a — good enough — explicit bound on the order of the G-function (that is, the rank of the $\mathbf{Q}(x)$-module generated by G and its derivatives), which — in a given situation — contradicts the structure of some explicit approximation function $P(x) = B(x) - \eta A(x)$. When this is achieved, the ultimate contradiction is in the supposition that $P(x) \in \mathbf{Q}[\![x]\!]$, that is that $\eta \in \mathbf{Q}$.

These arithmetic holonomy bounds are ultimately the main concern of this paper, and we take up a detailed introduction to them in our next section § 2.

Remark 1.2.5. In a very special case, a hint in this direction has been previously proposed (although without any application to a new irrationality proof) by Zudilin [Zud17], who isolated a condition on the linear forms $c_n = b_n - \eta a_n$ which implies an analytic continuation of the generating function $P(x) = \sum_{n=0}^{\infty} c_n x^n$ to a slit plane $\mathbf{C} \setminus [\beta, \infty)$. Whereas Apéry's use of the convergence radius focused on the decay rate $\beta^{-n+o(n)}$ of the coefficients c_n, Zudilin highlights the improved decay rate $(4\beta)^{-n^2+o(n^2)}$ of the sequence of *Hankel determinants* $\det(c_{i+j})_{i,j=0}^n$; this is indeed a well-known consequence [Pól28, Pom69] of the analyticity of $P(x)$ on $\mathbf{C} \setminus [\beta, \infty)$. The latter is in fact closely linked to the proof of Theorem 1.2.4 in this particular case of $\Omega = \mathbf{C} \setminus [\beta, \infty)$ with $\beta > 1/4$: if all $c_n \in \mathbf{Z}$, the Hankel determinants are also rational integers, therefore they vanish from some point onward if they decay at an exponential rate smaller than one, and finally this means $P(x) \in \mathbf{Q}(x)$ by the rationality criterion of Kronecker. Quantifying the denominator of the Hankel determinant in the case $c_n \in \mathbf{Q}$ leads as well to Zudilin's determinantal criterion for $\eta \notin \mathbf{Q}$. See § 2.7.7 for a precise formulation and a generalization. △

Remark 1.2.6 (A remark on exposition)**.** We take the point of view that the readership of this paper might include mathematicians not familiar with either the details of our previous paper [CDT21] or the methods of Diophantine analysis more broadly. At the risk of interrupting the flow of the exposition, we have included a number of expositional asides denoted by "basic remarks" throughout the paper which are intended to help orient the reader less familiar with this material; the expert should feel free to skip over these.

Remark 1.2.7 (A remark on notation)**.** We shall use $\mathbf{N} = \{0, 1, 2, \ldots\}$ to denote the natural numbers with zero, and $\mathbf{N}_{>0}$ to denote the positive integers. Depending on the context, $\mathbf{P}^1$ will signify either a scheme isomorphic to $\operatorname{Proj}\mathbf{Z}[T_0,T_1]$ (the projective line over $\mathbf{Z}$), or the complex manifold $\mathbf{P}^1(\mathbf{C})$ of its $\mathbf{C}$-valued points (the Riemann sphere with coordinate $z := T_0/T_1$, elsewhere commonly denoted $\widehat{\mathbf{C}}$ or $\mathbf{CP}^1$). A similar ambiguity is adopted for the modular stacks $Y_0(2)$ and $Y(2)$. The complex disc $D(0,R)$ of radius $R \in (0,\infty]$ in the relevant coordinate (always clear by the context, but most frequently denoted z) will be denoted by $\mathbf{D}_R$, and we shall write $\mathbf{D} := \mathbf{D}_1$ for the unit radius disc and $\overline{\mathbf{D}}$ for its closure in $\mathbf{C}$. The unit circle $\partial\overline{\mathbf{D}} = \{e^{2\pi i\theta} : \theta \in [0,1]\}$ is denoted $\mathbf{T}$ and its uniform measure $d\theta$ is denoted μ_{Haar}. For a connected complex manifold M, we shall denote by $\mathcal{O}(M)$ and $\mathcal{M}(M)$, respectively, the ring of holomorphic functions the field of meromorphic functions on M. The notation $\mathcal{O}(\overline{\mathbf{D}})$ and $\mathcal{M}(\overline{\mathbf{D}})$ is used for the corresponding functions on some unspecified open neighborhood of the closed unit disc $\overline{\mathbf{D}} \subset \mathbf{C}$. Throughout our paper, we will usually write $q := e^{\pi i\tau}$ for τ belonging to the upper half plane $\mathbf{H}$, although we will occasionally write $q = e^{2\pi i\tau}$. (As noted in [CDT21], this is forced upon us by historical convention, but we always use the first choice unless explicitly stated otherwise.) By a mild and harmless notational abuse, the *modular lambda function*

$$\lambda(q) := \frac{\left(\sum_{n\in 1+2\mathbf{Z}} q^{n^2/4}\right)^4}{\left(\sum_{n\in 2\mathbf{Z}} q^{n^2/4}\right)^4} = 16q\prod_{n=1}^{\infty}\left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8 : \mathbf{H} \to \mathbf{C}\smallsetminus\{0,1\},\ \{q=0\}\mapsto 0 \tag{1.2.8}$$

will be written in the cusp-filling coordinate $q \in \mathbf{D} := \{|q| < 1\}$ rather than $\tau = \log(q)/(\pi i)$. The letter e is generally reserved for the Euler constant $e \approx 2.718281$. Finally, we admit a minor notational abuse by adopting the convention of writing $X \smallsetminus \{A,B\} := X \smallsetminus (A\cup B)$ for any subsets A and B of a set X.

1.3. **The paths to Theorems A and C, and an outline of the paper.** The following leitfaden (Figure 1.3.0) gives in summary the logical structure of our paper. Here the pair of dotted lines indicates that there are two alternate paths to Theorems A and C, either through § 6 (by multivariable methods, based on measure concentration) or § 7 (by single variable methods, based on some Arakelov theory and Bost's inequality on evaluation heights). We also omit § B, which is most closely related (though there is no dependency in either direction) to § 7. There could be some (modest) economy if we restricted ourselves to the shortest possible proof of Theorem A. However, with a view to both future developments and applications, we felt it was better to include all these new ideas. In many ways this reflects our experience with our previous paper [CDT21] which included three proofs of the main holonomicity theorem [CDT21, Theorem 2.0.1]. One of the referees of that paper recommended removing one particular proof of this theorem whose ideas subsequently proved essential for the advances in this paper.

We now very briefly outline the paper. Section § 2 is mainly introductory, although §§ 2.5–2.11 present a basic form of our main results (which will not be proved until § 6) together with some applications, and §§ 2.12–2.13 outlines our approach to proving holonomy bounds, which is followed up in precise detail in § 3,

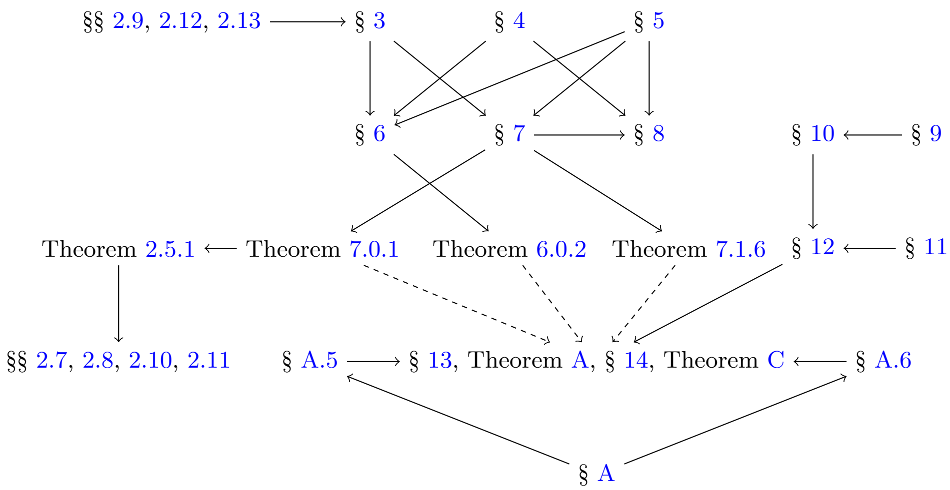

FIGURE 1.3.0. Leitfaden: paths to Theorems A and C

with elements of functional transcendence theory. In § 3, we also include some further exposition of related material in its proper historical context, intended to help place our ideas into a broader context. In § 4, we collect some basic facts concerning large deviations and the concentration of measure phenomenon in high dimensions. In § 5, we introduce the idea (possibly counterintuitive in light of the discussions in § 2) that it can sometimes be useful to integrate our putative functions despite introducing new denominators. Also included are some technical computations related to extra denominators arising from integrations, which follow from the prime number theorem. In § 6, we prove our first main holonomy bound Theorem 6.0.2. In § 7, based on the work of Bost and Charles [BC22], we prove our second main holonomy bound (or more precisely, several closely related bounds) using Bost's slopes method framework. In particular, Theorem 7.0.1 is essentially the bound of Bost–Charles in [BC22] incorporated with our treatment of denominators in § 6; Theorem 7.1.6 is a further improvement of Theorem 7.0.1 using the convexity property of a growth characteristic function which is closely related to the Bost–Charles bound and behaves similarly to a Nevanlinna characteristic function. In § 8, we unify our methods from §§ 6–7 and obtain, with an eye to future applications, the sharpest holonomy bound in our paper. In sections §§ 9–12 we return to a discussion of specific *templates* (situations in which the denominator types and singularities are fixed) in order to prepare for the application of our holonomy bounds to our main irrationality results. In § 9, we use the map of modular curves $Y(2) \to Y_0(2)$ to relate two templates over $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ and $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$, respectively. In § 10, we discuss some G-functions on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ with simple denominator types (most of them well-known, but also one which was surprising to us), and in § 11 we introduce certain local systems arising in [Zag09] which, contingent on a hypothetical linear dependence of 1, π^2, and $L(2,\chi_{-3})$, give rise to more G-functions. § 12 is concerned with proving the linear independence of all these functions over $\mathbf{C}(x)$. In §§ 13 and 14 we give the proofs of Theorems A and C respectively, using some explicit computations which are explained in detail in § A. (For a proof of Theorem C only, a number of subsections, including all of § 11, can also be omitted.) Finally, the short § B is intended as a showcase of the basic proof scheme, and can serve most particularly as an introduction to § 7.

2. The main arithmetic holonomy bound

We begin with a discussion of the dimension bounds in their simplest case.

2.1. **The algebraic case.** Our solution [CDT21] of the "unbounded denominators" conjecture was based on the following dimension upper estimate on a certain $\mathbf{Q}(x)$-linear space of algebraic functions. We called this type of result an *arithmetic holonomy bound*, and while our reason for this name remained obscure in [CDT21], we hope it should be vindicated by our present paper where we treat more general holonomic functions whose analytic continuations generate an infinite and non-solvable monodromy group. Given a holomorphic mapping $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$ on some neighborhood of the closed unit disc $\overline{\mathbf{D}} \subset \mathbf{C}$ and taking $\varphi(0) = 0$ with $|\varphi'(0)| > 1$, we established [CDT21, Theorem 2.1] the dimension upper bound

$$\dim_{\mathbf{Q}(x)} \mathcal{H}(\varphi) \leq e \cdot \frac{\displaystyle\int_{\mathbf{T}} \log^+ |\varphi| \, \mu_{\mathrm{Haar}}}{\log |\varphi'(0)|} \tag{2.1.1}$$

on the $\mathbf{Q}(x)$-linear span $\mathcal{H}(\varphi)$ of the $\mathbf{Z}[\![x]\!]$ formal power series $f(x)$ whose pullback $f(\varphi(z))$ also converges on a neighborhood of $\overline{\mathbf{D}}$. Here $\mathbf{T}$ is the unit circle, $\log^+ |x|$ is defined to be $\max(0, \log |x|)$, and μ_{Haar} is just the usual Haar measure on $\mathbf{T}$, so that the integral in the numerator can equally be written as $\int_0^1 \max\big(0, \log |\varphi(e^{2\pi i t})|\big)\, dt$.

Basic Remark 2.1.2. Suppose that $f(x)$ is a power series which extends to a holomorphic function on a domain $\Omega \subset \mathbf{C}$ containing the origin, of conformal mapping radius $\rho(\Omega, 0) > 1$. (See the beginning of § 2.7 for a precise definition of conformal mapping radius.) By definition, there consequently exists a biholomorphic map $\varphi : \mathbf{D} \to \Omega$ with $\varphi(0) = 0$ and $|\varphi'(0)| = \rho(\Omega, 0) > 1$. In turn, the holomorphy of $f(x)$ on Ω means exactly that the pulled back power series $f(\varphi(z)) \in \mathbf{C}[\![z]\!]$ converges on $\mathbf{D}$. The bound (2.1.1) then implies that the $\mathbf{Q}(x)$-vector space generated by $f(x) \in \mathbf{Z}[\![x]\!]$ *and its powers* is finite dimensional (since the powers of $f(x)$ also lie in $\mathcal{H}(\varphi)$), and thus $f(x)$ is algebraic (of some explicitly bounded degree). However, under these assumptions, one can already deduce the *rationality* of $f(x)$ from the Borel–Pólya Theorem 1.2.4. So the bound (2.1.1) is more interesting when φ is *not* univalent. (We will eventually find that the Borel–Pólya theorem too will be completely subsumed into holonomy bounds finer than (2.1.1), such as the bound (2.2.4) below, which is due to Bost and Charles [BC22], and ultimately our main new holonomy bound (2.5.4) in this paper.)

A non-univalent example is as follows. Suppose that $f(x) \in \mathbf{Z}[\![x]\!]$ can be analytically continued on any path from 0 in $\mathbf{C}$ avoiding both 0 and some fixed real number $\alpha > 0$. For example, take $f(x) = (1 - 4x)^{-1/2} = \sum \binom{2n}{n} x^n$, and $\alpha = 1/4$. Then one can take φ to be any holomorphic function with $\varphi(0) = 0$ but which has no other preimages of either 0 or α. One such function is

$$\varphi(z) = \alpha \lambda(z)$$

where λ is the modular λ function as given in (1.2.8). In this case, we have $|\varphi'(0)| = 16\alpha$. Hence, if $\alpha > 1/16$, we deduce that $f(x)$ is algebraic (with some degree explicitly bounded by (2.1.1) over $\mathbf{Q}(x)$). This example is already due to André.

The paper [CDT21] is concerned with the case when $\alpha = 1/16$, where there are infinitely many $\mathbf{Q}(x)$-linearly independent algebraic examples including

$$f(x) = \sum_{n=0}^{\infty} \binom{4n}{2n} x^n = \sqrt{\frac{1+\sqrt{1-16x}}{2-32x}};$$

but also the algebraicity fails without any additional hypothesis, as can be seen from the hypergeometric example $f(x) = \sum_{n=0}^{\infty} \binom{2n}{n}^2 x^n$, a case used in [And96] and further discussed in [And04, Appendix A]. We refer any further discussion of the algebraic case to [CDT21]. △

2.2. **Denominators.** For linear independence proofs, as suggested by the examples in § 1, we need holonomy bounds on functions in $\mathbf{Q}[\![x]\!]$ rather than $\mathbf{Z}[\![x]\!]$. Indeed, the holonomic coefficients of interest — such as $\eta = L(2, \chi_{-3})$ as our primary focus here — are conjecturally transcendental, and so any realization as numbers in a period matrix must necessarily involve a local system with an infinite global monodromy group. On the other hand, if $P(x) \in \mathbf{Z}[\![x]\!]$ lies in a holonomic module $\mathcal{H}(\varphi)$ attached to some $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$ with $\varphi(0) = 0$ and $|\varphi'(0)| > 1$, then (as noted previously) it would follow that $P(x)$ is algebraic. The *Grothendieck–Katz p-curvature conjecture* [Kat72, And04] (proved by Katz in many of the cases that are of geometric origin) informally equates the infinitude of the global monodromy group of an integrable connection with the nonvanishing of the *p-curvature operator* — the local obstruction to integrability modulo p — for a positive density of the primes p. But we remind the reader that, even for an irreducible linear homogeneous ODE, a single $\mathbf{Z}[1/S][\![x]\!]$ solution does not imply vanishing of the p-curvatures; rather, a *basis* of $\mathbf{Z}[1/S][\![x^{1/h}]\!]$ solutions does. As an example, the function $A(x) \in \mathbf{Z}[\![x]\!]$ in Apéry's argument (discussed in § 1) has integral coefficients but is not algebraic. To square this example with the remarks about $P(x)$ above, remember that the holonomy bounds are never being applied to $A(x) \in \mathbf{Z}[\![x]\!]$ itself, but rather to a (supposed for the contradiction!) $\mathbf{Q}$-linear combination $P(x) = B(x) - \eta A(x)$ of $A(x)$ and some other solution $B(x) \in \mathbf{Q}[\![x]\!]$ of the same ODE. This second solution $B(x)$ does indeed have denominators involving infinitely many primes.

From [FR17], we have a conjectural[4] understanding of the denominator types of Taylor series $P(x) = \sum a_n x^n \in \mathbf{Q}[\![x]\!]$ arising from G-functions: there should exist $A \in \mathbf{N}_{>0}$, $b \in \mathbf{Q}_{>0}$, and $\sigma \in \mathbf{N}$ such that

$$a_n A^{n+1} [1, \ldots, bn]^{\sigma} \in \mathbf{Z} \qquad \forall n \in \mathbf{N}; \tag{2.2.1}$$

here and throughout our paper (as noted previously in the introduction), $[1, \ldots, n]$ is used to denote the lowest common multiple of the first $\lfloor n \rfloor$ positive integers.

The most basic example is the G-function $\log(1-x)$. It has the type (2.2.1) with $A = 1$, $b = 1$, and $\sigma = 1$, but that form can in this case clearly be improved: only an n is needed out of the $[1, \ldots, n]$ clearance, in reflection of the fact that

$$\log(1-x) = \int \frac{dx}{x-1}$$

[4]This is an unconditional theorem for the case of G-functions that "arise from geometry," based on the existence of an F-crystal structure at all but the finitely many primes of bad reduction, cf. [And89, § V app.]. One can also be more precise: if $\mathcal{L}(f) = 0$ for some nonzero r^{th}-order Fuchsian operator having for $x = 0$ local exponents rational numbers with denominators dividing b, then the denominators form of f may be taken as $A^{n+1}[1, \ldots, bn + b_0]^{r-1}$ for some $A \in \mathbf{N}_{>0}$ and $b_0 \in \mathbf{Z}$.

is an integral of an algebraic function. It turns out, cf. § 5, that it will ultimately be important to exploit such refinements from integrals. In any case, the necessity of at least the $[1,\dots,bn]$ denominators forces us to venture outside of the proper[5] scope of the theory of formal-analytic arithmetic surfaces [Bos20, BC22].

With the presence of denominators, for given holomorphic $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$ and parameters $b \in \mathbf{Q}_{\geq 0}$ and $\sigma \in \mathbf{N}$, we define the *holonomic module* $\mathcal{H}(\varphi; b; \sigma)$ to be the $\mathbf{Q}(x)$-linear span of all the formal functions of the form

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\dots,bn]^{\sigma}} \in \mathbf{Q}[\![x]\!], \qquad a_n \in \mathbf{Z} \quad \forall n \in \mathbf{N} \tag{2.2.2}$$

whose φ-pullbacks $f(\varphi(z))$ converge on $\mathbf{D}$. The proofs in [CDT21, § 2] extend routinely to establish a first result in this direction:

$$\dim_{\mathbf{Q}(x)} \mathcal{H}(\varphi; b; \sigma) \leq e \cdot \frac{\int_{\mathbf{T}} \log^+ |\varphi| \, \mu_{\mathrm{Haar}}}{\log|\varphi'(0)| - \tau}, \tag{2.2.3}$$

where $\tau := b\sigma$ and we now assume that φ has the *conformal size* $|\varphi'(0)| > e^{\tau}$.

Unfortunately, the holonomy bound (2.2.3), which worked nicely in the asymptotic framework of [CDT21] where the absolute numerical coefficient was immaterial, is now far too crude to prove the irrationality of $L(2,\chi_{-3})$. It is then of interest to know the least possible value that may take the place of the constant e in the bound (2.2.3). Progress was made by Bost and Charles [BC22, Corollary 8.3.5] who, in the original $\sigma = 0$ case of [CDT21], established the finer bound by

$$\frac{\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)| \, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)|}. \tag{2.2.4}$$

In 2023, in response to our question about a similar dimension bound for the general holonomic modules $\mathcal{H}(\varphi; b; \sigma)$, Charles explained to us how their proof can be directly generalized to obtain

$$\dim_{\mathbf{Q}(x)} \mathcal{H}(\varphi; b; \sigma) \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)| \, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - b\sigma}. \tag{2.2.5}$$

This in particular implies (see, for example, Corollary 8.1.14) that the coefficient e in (2.1.1) and (2.2.3) can be taken down to the better constant 2. Bost and Charles's work has been a major stimulus for our exploration of the applications to irrationality. Inspired by [BC22], but going outside of their framework of formal-analytic arithmetic surfaces and incorporating an idea of Perelli and Zannier [PZ84], we prove in § B the reduction $e \rightsquigarrow 2$ in (2.2.3). In § 7, we carry this further based on some of Bost and Charles's results from [BC22] re-interpreted for analytic purposes into Bost's prior method of evaluation heights, in order to generalize (2.2.4) and (2.2.5) to incorporate a refined denominator term; see Theorem 2.5.1 for a special case of our bounds from § 7. Our companion treatise [CDT24] of the irrationality of the 2-adic zeta value $\zeta_2(5)$ explores these bounds in a wider context.

[5]A natural framework would be the construction and comparison of integrable connections over formal-analytic arithmetic varieties and their algebraizations. We do not attempt to get into such a concept in the present paper, apart from raising one specific finiteness problem in § 15.2, but we do hope to turn to it on another occasion.

2.3. **A preview of the various holonomy bounds.** The core of our present paper consists of refined holonomy bounds that improve (2.2.5) and unify the proof methods behind (2.2.3) and (2.2.4). One aspect of these bounds is to improve the $\tau = b\sigma$ term in (2.2.3); here the high-dimensional methods ultimately yield a more precise information, although the difference is invisible to all our applications in this paper. The other aspect is to carry out a more refined complex analytic estimate (see § 2.13.10 and § 2.13.13 for a summary of ideas) to further improve the double integral in (2.2.4); here the improvements are the same in the single variable as in the high-dimensional treatments. One technical novelty is a probabilistic input from large deviations theory which accommodates the $e \rightsquigarrow 2$ lowering in (2.1.1) even in the elementary multivariable framework of our original analysis in [CDT21, § 2]. This is established through a Diophantine approximation argument in d auxiliary variables, and the point of achieving the $e \rightsquigarrow 2$ coefficient improvement in precisely this way is that the high-dimensional geometric features of the $d \to \infty$ asymptotic make an additional room for further independent improvements. The sharpest holonomy bound (Theorem 8.0.1) that we have in this paper is a product of the measure concentration feature in the high-dimensional evaluation module.

For the applications in this paper, including Theorems A and C, the finest improvement concerning the general denominators does not make a difference. We have two general simplified lines to these theorems. One is via Theorem 6.0.2 using the high-dimensional techniques in a basic Siegel lemma framework, but another is via Theorem 7.0.1 and alternatively Theorem 7.1.13 using single variable methods. (See § 1.3 for more details on the dependencies between different sections of this paper, and the various paths to Theorems A and C.) For the application to Theorem A, Theorem 7.0.1 gives the weakest passable bound (sufficient by only a narrow margin) compared to these two other theorems; while its "convexity refinement," Theorem 7.1.10, gives a stronger bound than either of them. The proof of Theorem 7.0.1 is a direct combination of the work of Bost and Charles, together with our improvement of the $\tau = b\sigma$ term in a relatively simple setting (see Theorem 2.5.1; this simple setting allows us to get the optimal improvement of τ even without a high dimensional method), and a computation in § 5 to accommodate added powers of n in the denominator types (7.0.1).

To get the stronger bounds that handle Theorem A by a more comfortable margin, we use more refined complex analytic estimates to prove Theorems 6.0.2 and 7.1.6. In the case of the former, the large deviations input is used not only to reach the improvement of the denominators rate term τ, but also to obtain a replacement of the Bost–Charles double integral by a more elementary *rearrangement integral*[6] which we introduce in § 2.4; the proofs here are fully independent of [BC22]. On the other hand, based on [BC22] and Theorem 7.0.1, we undertake a closer study of the optimal archimedean estimates for the heights of the evaluation maps in Bost's slopes framework, and employ these improvements to prove Theorem 7.1.6. This is what we dub the *improvement from convexity*, a choice of terminology that refers to a classical theorem in the value distribution theory

[6]A *rearrangement integral* here refers to a more general set of functions than the $\log|\varphi|$ in the integrand of $\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t})|)^* \, dt$. The latter, as we will see in § 8.1, is larger than the Bost–Charles double integral. In general, we replace the φ inside this integrand by a piecewise weighted combination of the functions $z \mapsto \varphi(rz)$, using a suitable set of radii r that facilitate our refined complex analytic estimate.

of meromorphic functions: the Nevanlinna characteristic $T(r,\varphi)$ of a meromorphic function φ is a convex increasing function of $\log r$. Further, if we choose a certain heuristically optimal Hermitian structure on the evaluation module of auxiliary polynomials, the argument of Theorem 7.1.6 leads to a heuristically optimal bound which we formulate as Theorem 7.6.4, still using single variable methods. In the basic denominators capping such as we introduce already in Theorem 2.5.1 further down in this introduction, Theorem 7.6.4 is the same as Theorem 8.0.1 (cf. Remark 8.0.6), and we expect (cf. Remark 7.6.7) both to give a stronger bound than Theorem 6.0.2.

We proceed now to describe some of these basic improvements, and then state a first form of our new holonomy bounds.

2.4. **Variants of the Nevanlinna growth characteristic.** From the starting bound (2.2.3), on further pursuing [CDT21, Remark 2.3.3], the multiple variables naturally improve Nevanlinna's growth characteristic term $\int_0^1 \log^+ |\varphi(e^{2\pi i t})|\, dt$ to the manifestly smaller *rearrangement integral* $\int_0^1 t \cdot (\log|\varphi(e^{2\pi i t})|)^*\, dt$; here and throughout our paper, we follow the classical analysis custom to designate by

$$g^*(t) := \inf_{s\in\mathbf{R}} \{\mathbf{P}\,(x\in(0,1)\,:\,g(x)>s)\le t\} = \inf_{s\in\mathbf{R}} \left\{\int_0^1 \chi_{g^{-1}([s,\infty))}\,dt \le t\right\} \tag{2.4.1}$$

the *increasing rearrangement* of a measurable function $g:(0,1)\to\mathbf{R}$. (See Basic Remark 2.4.4.) This is the unique[7] nondecreasing measurable function that has the same distribution function as g. We thus have

$$\int_0^1 2t\cdot g^*(t)\,dt = \int_0^1\int_0^1 \max(g(s),g(t))\,ds\,dt, \tag{2.4.2}$$

inviting a comparison to the Bost–Charles double integral term from (2.2.4). We will see in § 8.1 that the latter is always, and in practice only slightly, smaller than the former.

It is, however, the left-hand side of (2.4.2) that arises naturally in the probabilistic character of our new argument. For our discussion here it suffices to note the trivial inequality

$$\int_0^1 2t{\cdot}g^*(t)\,dt \le \int_0^1 2\max(g^*(t),0)\,dt = \int_0^1 2\max(g(t),0)^*\,dt = 2\int_0^1 \max(g(t),0)\,dt \tag{2.4.3}$$

for any measurable function g, and so in particular this recovers the $e \rightsquigarrow 2$ coefficient reduction from a genuinely high-dimensional perspective following [CDT21, § 2] which is in some sense an approach "orthogonal" to the single variable analyses of either [BC22] or § 7. (These latter approaches have an Arakelovian character, and carry their own and different refinement of the Nevanlinna growth characteristic, which in § 7.1.1 we dub the *Bost–Charles characteristic*.) Now the point is that the $d\to\infty$ argument further allows for an analogous "denominator increasing rearrangements" improvement of the term $\tau = b\sigma$ in the extension (2.2.3) to $\mathbf{Q}[\![x]\!]$ functions. Some such improvement is essential for all our proofs of Theorems A and C. We also do give a single variable treatment in § 7 of the main results of § 6.

[7]Up to functions vanishing outside of a set of measure zero. Some authors prefer to use the term *nondecreasing rearrangement function*.

The high dimensional method, on the other hand, leads to an even more precise bound in the denominators aspect, a refinement that could be useful in further developments or applications of our method.

Basic Remark 2.4.4. A basic way to understand the definition of $g^*(t)$ in equation (2.4.1) is as follows. Assume that $g(t)$ is a continuous (and hence bounded) function on $[0,1]$. If $g(t)$ is monotonically increasing, then $g^*(t) = g(t)$. For $g(t)$ arbitrary, let $g_n(t)$ for $n \geq 1$ denote the piecewise constant step function which takes the value $g(k/n)$ on the interval $I_k = [(k-1)/n, k/n)$ for $k = 1, \ldots, n$ (extending the final interval I_n to include 1). The functions $g_n(t)$ converge uniformly to $g(t)$ as $n \to \infty$. Now let $g_n^*(t)$ denote the step function which is also constant in the n intervals $I_1, \ldots, I_n$, except now taking the n respective values

$$\{g(1/n), g(2/n), g(3/n), \ldots g(n/n)\}$$

rearranged in *increasing order* (hence the name). Then $g_n^*(t)$ is the increasing rearrangement of $g_n(t)$, and the functions $g_n^*(t)$ converge uniformly to $g^*(t)$. △

2.5. **Arithmetic holonomy bounds, basic form.** Our first main result is the following simultaneous strengthening of all the *holonomy bounds* or arithmetic rationality or algebraicity criteria that we have explicitly stated so far.

Theorem 2.5.1. *Consider two positive integers $m, r \in \mathbf{N}_{>0}$ and an $m \times r$ rectangular array of nonnegative real numbers $\mathbf{b} := \left(b_{i,j}\right)_{1 \leq i \leq m,\, 1 \leq j \leq r}$, all of whose columns are of the form:*

$$0 = b_{1,j} = \cdots = b_{u_j,j} < b_{u_j+1,j} = \cdots = b_{m,j} =: b_j, \qquad \forall j = 1, \ldots, r,$$

for some $u_j \in \{0, 1, \ldots, m\}$. Let

$$\sigma_i := b_{i,1} + \ldots + b_{i,r}, \qquad i = 1, \ldots, m$$

be the i-th row sum, and define

$$\tau(\mathbf{b}) := \frac{1}{m^2} \sum_{i=1}^{m} (2i-1)\sigma_i = \sigma_m - \frac{1}{m^2} \sum_{j=1}^{r} u_j^2 b_j \in [0, \sigma_m]. \tag{2.5.2}$$

Further, consider a holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ with derivative (conformal size) satisfying $|\varphi'(0)| > e^{\sigma_m}$.

Suppose there exists an m-tuple $f_1, \ldots, f_m \in \mathbf{Q}[\![x]\!]$ of $\mathbf{Q}(x)$-linearly independent formal functions with denominator types of the form

$$f_i(x) = \sum_{n=0}^{\infty} a_{i,n} \frac{x^n}{[1, \ldots, b_{i,1} \cdot n] \cdots [1, \ldots, b_{i,r} \cdot n]}, \qquad a_{i,n} \in \mathbf{Z}, \tag{2.5.3}$$

such that $f_i(\varphi(z)) \in \mathbf{C}[\![z]\!]$ is the germ of a meromorphic function on $|z| < 1$, for all $i = 1, \ldots, m$. Then we have the bound

$$m \leq \frac{\displaystyle\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)| \, \mu_{\mathrm{Haar}}(z) \mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - \tau(\mathbf{b})}. \tag{2.5.4}$$

If, moreover, all functions f_i are a priori *assumed to be holonomic, the condition $|\varphi'(0)| > e^{\sigma_m}$ can be relaxed to $|\varphi'(0)| > e^{\tau(\mathbf{b})}$.*

With elementary methods based on the phenomenon of measure concentration in high dimensions, we prove directly in § 6 the following variant using the increasing rearrangement function:

$$m \leq \frac{\iint_{\mathbf{T}^2} \log\left(\max(|\varphi(z)|, \varphi(w)|\right) \, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - \tau(\mathbf{b})} = \frac{\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t}))^* \, dt}{\log|\varphi'(0)| - \tau(\mathbf{b})}. \tag{2.5.5}$$

To highlight the similarity of $\tau(\mathbf{b})$ with the increasing rearrangement function that emerges from a simple probabilistic consideration, let us note (writing $\sigma_0 := 0$) that the weighted average in (2.5.2) can be expressed in a form rather similar to (2.4.2):

$$\tau(\mathbf{b}) = \sum_{i=1}^{m} \sigma_i \int_{(i-1)/m}^{i/m} 2t\, dt = \int_0^1 2t \cdot \left(\sum_{i=1}^{m} (\sigma_i - \sigma_{i-1}) \chi_{[0,i/m]}(t) \right) dt. \tag{2.5.6}$$

Noting the monotonicity of the step function $\sum_{i=1}^{m}(\sigma_i - \sigma_{i-1})\chi_{[0,i/m]}(t)$, our requirement that $\mathbf{b}$ is column-wise nondecreasing serves as the counterpart for denominators of the increasing rearrangement function $(\log|\varphi|)^*$. As remarked above, we will see in § 8.1 that the Bost–Charles integral in (2.5.4) can be tightly majorized by $\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t})|)^* \, dt$, with the effect that the bound (2.5.4) implies the bound (2.5.5). For either of the rearrangement integrals (2.5.5) and (2.5.6), the integration weight $2t$ arises as the cumulative distribution function of $([0,1], \mu_{\mathrm{Lebesgue}})$. One mechanism for both these improvements over (2.2.3) is held by the concentration of measure phenomenon § 4; it is explained in § 6.1.

For the rather rudimentary shape of the denominator type form (2.5.3) in our statement of Theorem 2.5.1, a single variable proof is nevertheless also possible, as we discover with the slopes method in § 7.3. In that context, both the denominators rate term $\tau(\mathbf{b})$ and the Bost–Charles double integral term in (2.5.4) emerge from the computation of the covolume of the Euclidean lattice of auxiliary polynomial functions chosen in the usual Diophantine analysis proof scheme: the former as the minimizer of a multivariable quadratic form arising from a basic template sought for the integral structure, and the latter as the "infinite part" based on a combination (due to Bost and Charles [BC22, § 5]) of the Poincaré–Lelong formula in complex analysis and the arithmetic Hilbert–Samuel formula in Arakelov theory.

2.6. **Siegel's G-functions.** As discussed above, Theorem 2.5.1 has a crude qualitative corollary which we may read as an arithmetic holonomicity criterion. It is due to André [And89, § VIII 1.6] (where the set of places V in *loc. cit.* must be assumed to be finite); in a slightly different context, the first holonomicity result of such a kind is probably the one discovered by Perelli and Zannier [PZ84, Thm. 1 B].

Corollary 2.6.1. *If a formal function $f \in \mathbf{Q}[\![x]\!]$ has rational coefficients of the form*

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\ldots,b_1 n]\cdots[1,\ldots,b_r n]}, \qquad a_n \in \mathbf{Z} \tag{2.6.2}$$

and admits an analytic mapping $\varphi : (\mathbf{D},0) \to (\mathbf{C},0)$ with conformal size $|\varphi'(0)| > e^{b_1+\ldots+b_r}$ and such that the composite function germ $f(\varphi(z)) \in \mathbf{C}[\![z]\!]$ is the germ of *a meromorphic function on* $\mathbf{D}$*, then* $f(x)$ *is a* holonomic function*: there exists a nonzero linear differential operator* $\mathcal{L}$ *with* $\mathbf{Q}[x]$*-coefficients that satisfies* $\mathcal{L}(f) = 0$.

In this paper, we will exhibit and exploit such $f \in \mathbf{Q}[\![x]\!]$ whose holonomicity can be recognized by this criterion. The special form of the denominators (2.6.2) then situates us more specifically into the context of Siegel's theory of G-functions; in particular, see Remark 3.2.12 for a discussion, the linear differential operator $\mathcal{L}$ can *a posteriori* be taken to be of the Fuchsian class with only regular singular points and with rational exponents [DGS94, III 6.1, VII 2.1, and VIII 1.5]. A major open question, which is closely related to the discussion of § 15.2 with implications to irrationality proofs and effective Siegel integral points problems, is to control the possible *apparent* singularities of the linear differential operator $\mathcal{L}$ in a minimal-order *inhomogeneous* ODE $\mathcal{L}(f) \in \mathbf{Q}[x]$.

Basic Remark 2.6.3. A simplest example is $f(x) = \log(1-x)$, with type given by $(b_1, \ldots, b_r) = (1)$ and minimal differential operator $\mathcal{L} := (1-x)(d/dx)^2 - (d/dx)$, varying holonomically on the domain $\Omega = \mathbf{C} \smallsetminus \{1\}$ to define a rank-2 local system

$$\mathrm{Span}_{\mathbf{C}}\big\{1, \log(1-x)\big\}$$

on Ω with monodromy the infinite cyclic group generated by the unipotent matrix

$$T := \begin{pmatrix} 1 & 0 \\ -2\pi i & 1 \end{pmatrix}.$$

This expresses the fact that the analytic continuation process T — the *local monodromy operator* — for $\log(1-x)$ under the counterclockwise direction along a simple closed loop encircling the singularity $\{1\}$ leaves $f_1 := 1$ invariant but adds to $f_2 := \log(1-x)$ the *period* $-2\pi i$ times f_1:

$$T^k(\log(1-x)) = \log(1-x) - 2\pi i k, \qquad T^k(1) = 1.$$

This holonomic example is furthermore recognized as a case of the holonomicity criterion Corollary 2.6.1, for instance with the multivalent choice $\varphi(z) := 1 - e^{-Rz}$ for any $R > e$, or the multivalent choice $\varphi(z) := \lambda(z)$ with $|\varphi'(0)| = 16 > e$, or the univalent choice $\varphi(z) := 4z/(1+z)^2$ with $|\varphi'(0)| = 4 > e$.

In Theorem 2.5.1, the denominators type is captured by the 2×1 matrix $\mathbf{b} = (0,1)^{\mathrm{t}}$, with $\tau(\mathbf{b}) = (1 \cdot 0 + 3 \cdot 1)/2^2 = 3/4$. For the choice $\varphi(z) := 4z/(1+z)^2$, the holonomy quotient is $\log 4/(\log 4 - 3/4) \approx 2.1787$, an upper bound on the dimension $m = 2$ of this local system. △

In § 11, we will make a thorough study of Zagier's holonomic functions [Zag09] that endow the numbers $\zeta(2)$ and $L(2, \chi_{-3})$ similarly as periods in a much more complicated local system spread over the domain $\Omega = \mathbf{C} \smallsetminus \{0, 1/9, 1\} \cong \mathbf{H}/\Gamma_0(6)$. For this local system, which emerged from analyzing the form of the recursion from Apéry's $\zeta(2)$ irrationality proof [Ape79, Coh78, vdP79] and is based on the theory of Eichler integrals, we will now have the main integrality type $x^n/[1, \ldots, n]^2$. We will then reduce the $\mathbf{Q}$-linear independence problem of $1, \zeta(2), L(2, \chi_{-3})$ to a Diophantine analysis problem on the nonexistence of a G-function of the type $x^n/[1, \ldots, n]^2$ and with certain analytic properties: specifically, our task becomes to prove that Zagier's local system cannot contain a nonzero $\mathbf{Q}[\![x]\!]$ element which is regular — *overconvergent* — at the singularities $\{0, 1/9\}$.

A direct application of (2.5.4), see § 2.11 further down in this section, suffices for proving the irrationality of the mixed period

$$L(2,\chi_{-3}) - \pi\frac{\log 3}{3\sqrt{3}} = L(2,\chi_{-3}) - L(1,\chi_{-3})\log 3.$$

(Another irrationality result for a mixed period is Beukers's proof [Beu87, Thm 4] using modular forms that $\zeta(3) - 5\sqrt{5}\,L(3,\chi_5) \notin \mathbf{Q}(\sqrt{5})$.) For the irrationality proof of the pure $L(2,\chi_{-3})$, as discussed in § 2.3, we need an even finer result than this to also take into account the integrals of the functions. These more elaborate versions of Theorem 2.5.1 (including Theorems 6.0.2, 7.0.1, 7.1.6, and 7.1.13) are deferred to § 6 and § 7 below where they are proved. The particular application to Theorem A is fairly delicate, and among the many local systems generating $\zeta(2)$ and $L(2,\chi_{-3})$ among their holonomic coefficients, the choice that ends up working for us is highly reducible (although with nonsolvable monodromy) and involves integrations that lead to denominators essentially[8] of the form $n[1,\ldots,2n]^2$.

2.7. **Univalent holonomy bounds and an arithmetic characterization of the logarithm.** We now consider the specialization of Theorem 2.5.1 to the setting where the map φ is univalent. We remark that although, for general φ, we have various improvements of Theorem 2.5.1, such as Theorems 7.1.6 and 7.6.4 (assuming $\mathbf{e}$ in *loc. cit.* is $\mathbf{0}$), in the case of univalent φ, all these reduce to the same Theorem 2.7.1 below.

For $\Omega \subset \mathbf{C}$ a contractible domain containing 0, the Riemann mapping theorem supplies a biholomorphic map $\varphi : \mathbf{D} \xrightarrow{\simeq} \Omega$ with $\varphi(0) = 0$, which by Schwarz's lemma is uniquely defined up to pre-composing by a circle rotation. That makes the absolute value $|\varphi'(0)| \in (0,\infty]$ well-defined; we denote it by $\rho(\Omega,0)$ and call it *the conformal mapping radius of the pointed contractible domain* $(\Omega,0)$. The holomorphic mapping $\varphi : \mathbf{D} \to \mathbf{C}$ is said to be *univalent* if it is biholomorphic onto its image, or equivalently, if $\varphi : \mathbf{D} \hookrightarrow \mathbf{C}$ is injective.

Theorem 2.7.1 (Univalent holonomy bound)**.** *Under the notations and assumptions of Theorem 2.5.1, consider $\Omega \subset \mathbf{C}$ a contractible domain with $0 \in \Omega$ and having a conformal mapping radius $\rho(\Omega,0) > e^{\tau(\mathbf{b})}$. For any m-tuple of $\mathbf{Q}(x)$-linearly independent formal functions of the type (2.5.3) and meromorphic in Ω, the following holonomy bound holds:*

$$m \leq \frac{\log \rho(\Omega,0)}{\log \rho(\Omega,0) - \tau(\mathbf{b})}.$$

Proof. This follows directly from Theorem 2.5.1. The point to observe is that the Bost–Charles double integral term satisfies the inequality

$$\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) \geq \log|\varphi'(0)|,$$

with equality if and only if $\varphi : \mathbf{D} \hookrightarrow \mathbf{C}$ is univalent on the open disc. To see this, simply observe that the univalence is equivalent to having the bivariate holomorphic function

$$\frac{\varphi(z) - \varphi(w)}{z - w} = \varphi'(0) + O\left(|z| + |w|\right) \in \mathcal{O}(\mathbf{D}^2)$$

[8]More precisely, of the form $n[1,\ldots,2n+3]^2$, but this can more or less be treated as having the shape $n[1,\ldots,2n]^2$, by Remark 6.0.12.

to be nonvanishing throughout the unit polydisc. Hence the function

$$G(z,w) := \log \left| \frac{\varphi(z)-\varphi(w)}{z-w} \right| \; : \; \mathbf{D}^2 \to \mathbf{R} \cup \{-\infty\}$$

is plurisubharmonic, and harmonic if and only if φ is univalent. Both claims now follow upon remarking that $G(0,0) = \log|\varphi'(0)|$ while

$$\iint_{\mathbf{T}^2} G(z,w)\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) = \iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w),$$

by the basic integral $\iint_{\mathbf{T}^2} \log \frac{1}{|z-w|}\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) = 0$. □

We note the following application of Theorem 2.7.1 to the logarithm function, which is the example of Basic Remark 2.6.3.

Theorem 2.7.2. *Suppose $f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbf{Q}[\![x]\!]$ is a power series such that:*

(1) $[1,\ldots,n]a_n \in \mathbf{Z}$ for all $n \in \mathbf{N}$.

(2) $f(x)$ is holomorphic on $\mathbf{C} \smallsetminus [1,\infty)$.

Then

$$f(x) = Q_0(x) + Q_1(x)\log(1-x)$$

for some rational functions $Q_0, Q_1 \in \mathbf{Q}\left[x, \dfrac{1}{1-x}\right] \subset \mathbf{Q}(x)$.

We view Theorem 2.7.2 as an *arithmetic characterization* of the logarithm function.

Proof. We consider the contractible domain $\Omega := \mathbf{C} \smallsetminus [1,\infty)$, of conformal mapping radius $\rho(\Omega,0) = 4$ with the Riemann map $\varphi(z) = 4z/(1+z)^2$. Applying Theorem 2.7.1 with $m=3, r=1$, and $\mathbf{b} = (0,1,1)^{\mathrm{t}}$ with $\tau(\mathbf{b}) = 8/9$, the numerology

$$\frac{\log 4}{\log 4 - 8/9} = 2.787050\ldots < 3 \tag{2.7.3}$$

proves that there is no third such function $\mathbf{Q}(x)$-linearly independent from the two known examples $f_1 = 1$ and $f_2 = \log(1-x)$ for the type (2.6.2) with $(b_1,\ldots,b_r) = (1)$ and analytic on $\Omega = \mathbf{C} \smallsetminus [1,\infty)$. This means that all such examples are of the form $Q_0(x) + Q_1(x)\log(1-x)$ with $Q_0(x), Q_1(x) \in \mathbf{Q}(x)$.

At this point, we know that $f(x)$ is regular (holomorphic) on $\mathbf{C} \smallsetminus \big([1,\infty) \cap \overline{\mathbf{Q}}\big)$, and that every point $x \neq 1$ in $\mathbf{C}$ is at worst a meromorphic pole of $f(x)$. It remains to prove two things:

(i) $Q_0(x)$ and $Q_1(x)$ are from the subring $\mathbf{Q}\left[x, \frac{1}{x}, \frac{1}{1-x}\right]$ of $\mathbf{Q}(x)$.

(ii) It is impossible to have $Q_0(x), Q_1(x) \in \mathbf{Q}[x, 1/x]$ without having both $Q_0(x), Q_1(x) \in \mathbf{Q}[x]$.

Indeed, (ii) gives what we want assuming (i) and upon changing $f(x)$ to $(1-x)^k f(x)$ with a sufficiently high power $k \in \mathbf{N}$ to clear the $(1-x)$ denominators from $Q_0(x)$ and $Q_1(x)$.

We first prove (ii). Suppose $Q_0(x)$ and $Q_1(x)$ are not both in $\mathbf{Q}[x]$. If

$$Q_1(x) \in \mathbf{Q}[x] \subset (1/x)\mathbf{Q}[x],$$

then $Q_1(x)\log(1-x)$ is holomorphic at $x=0$, but then $Q_0(x) = f(x) - Q_1(x)\log(1-x)$ is also holomorphic at $x=0$ and then $Q_0(x) \in \mathbf{Q}[x]$. Hence we may assume that $Q_1(x) \in \mathbf{Q}[x,1/x] \smallsetminus \mathbf{Q}[x]$. After multiplying $f(x)$ by the correct power of x and a suitably divisible positive integer, we may assume that $Q_1(x) \in \big(\mathbf{Z}[x] + \mathbf{Z}\cdot x^{-1}\big) \smallsetminus \mathbf{Z}[x]$ and $Q_0(x)$ (which is now holomorphic by the argument above) lies in $\mathbf{Z}[x, 1/x]$ and hence also in $\mathbf{Z}[x]$, and still with the denominator property (1) in place. In turn, upon subtracting from $f(x)$ a suitable element of $\mathbf{Z}[x] + \log(1-x)\mathbf{Z}[x]$, we are left with analyzing the case

$$f(x) = \frac{q_1 \log(1-x)}{x}$$

with $q_1 \in \mathbf{Z} \smallsetminus 0$. But then the x^{p-1} coefficient of $f(x)$ is equal to q_1/p, which, when $p > |q_1|$ is a prime, is not of the required form (1). This completes the reduction step (ii).

We now consider (i). Suppose for contradiction that the rational functions $Q_0(x)$ and $Q_1(x)$ are not from the subring $\mathbf{Q}\left[x, \frac{1}{x}, \frac{1}{1-x}\right]$; then at least one of them will have a pole $\alpha \in \overline{\mathbf{Q}} \smallsetminus \{0,1\}$.

Fix a complex embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$, and consider firstly the case that $\alpha \notin (1,\infty)$ for at least one of the poles of $Q_0(x)$ or $Q_1(x)$. In that case, our assumption that $f(x)$ is holomorphic at α implies that

$$Q_0(x) = \frac{V(x)}{(x-\alpha)^k}, \quad Q_1(x) = \frac{-U(x)}{(x-\alpha)^k}$$

with some *positive* integer $k \in \mathbf{N}_{>0}$ and some rational functions $U, V \in \overline{\mathbf{Q}}(x)$ regular and nonzero at $x = \alpha$. Setting $x = \alpha$ in the equation

$$(x-\alpha)^k f(x) = V(x) - U(x)\log(1-x)$$

yields a nontrivial vanishing combination $V(\alpha) - U(\alpha)\log(1-\alpha) = 0$ with nonzero algebraic number coefficients $U(\alpha), V(\alpha) \in \overline{\mathbf{Q}}^{\times}$. But this contradicts the Hermite–Lindemann–Weierstrass theorem on transcendental values of the function $\log(1-x)$ on $\overline{\mathbf{Q}} \smallsetminus \{0,1\}$.

It remains to handle the case that *all* poles $\alpha \neq 0, 1, \infty$ of Q_0 and Q_1 belong to $\alpha \in (1,\infty) \cap \overline{\mathbf{Q}}$, and that this set of poles is nonempty. Here, our $f \in \mathcal{O}(\mathbf{C} \smallsetminus [1,\infty))$ holomorphy condition does not rule out a meromorphic pole at $x = \alpha$, and we need a different argument. As the set of poles in consideration is stable under $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, our assumption implies that all Galois conjugates of α lie in $(1,\infty) \cap \overline{\mathbf{Q}}$. We then deduce from the product formula that there is a prime p and a choice of a pole $\alpha \in \overline{\mathbf{Q}} \cap \mathbf{C}_p$ lying within the open disc $|x|_p < 1$. Then we get the same contradiction p-adically, upon citing[9] Mahler's theorem [Mah19a] on the transcendence of all convergent values of the p-adic exponential function at nonzero algebraic arguments; which is equivalent to the transcendence of all values of the p-adic logarithm function $\log(1-x)$ at the algebraic points of the punctured open unit disc $0 < |x|_p < 1$. □

Remark 2.7.4. Theorem 2.7.2 and its proof also holds with, for example, (1) relaxed to the form

$$\left[1, \ldots, \left(1 + \frac{1}{100}\right)n\right];$$

[9]The p-adic counterpart of the full Hermite–Lindemann–Weierstrass theorem on the algebraic independence of special values of the exponential function is a well-known and still-unresolved conjecture. We refer to Nesterenko's work [Nes08], for partial results, and [Nes19, § 2.4], for an overview of the subject and an introduction to Mahler's argument.

but then with the weaker conclusion $Q_0, Q_1 \in \mathbf{Q}[x, 1/x, 1/(1-x)]$ from step (i) alone, where indeed $1/x$ can no longer be removed, as instanced by the function $f(x) = 100! \cdot \log(1-x)/x$. Here the constant $1 + 1/100$ could equally be replaced by any element of $(1, (3\log 2)/2) = (1, 1.03972\ldots)$. △

Remark 2.7.5. In the conclusion of Theorem 2.7.2, we can completely characterize the possible $Q_i(x)$. Namely, $f(x) = Q_0(x) + Q_1(x)\log(1-x)$ has the required form if and only if the following two conditions hold:

(1) $Q_1(x) \in \mathbf{Z}[x, 1/(1-x)]$.
(2) $Q_0(x)$ lies in the $\mathbf{Z}[x, 1/(1-x)]$-module generated by $x^n/[1,\ldots,n]$ for each n — equivalently, generated by x^q/q for each prime power q.

This gives the full description by generators and relations of the (infinite) $\mathbf{Z}[x, 1/(1-x)]$-module of solutions in Theorem 2.7.2.

It is plain that these conditions yield the requirements of Theorem 2.7.2. To prove the converse, consider an $f(x) = Q_0(x)+Q_1(x)\log(1-x)$ in the theorem. The conclusion for $Q_0(x)$ is clear once we establish the conclusion for $Q_1(x)$. Without loss of generality (after multiplying by a power of $(1-x)$), it suffices to show that if $Q_i(x) \in \mathbf{Q}[x]$, then $Q_1(x) \in \mathbf{Z}[x]$. If $Q_1(x) \notin \mathbf{Z}[x]$, then there exists a prime p and a monomial $q_{1,m}x^m$ of $Q_1(x)$ such that the p-adic valuation $\mathrm{val}_p(q_{1,m}) < 0$ is negative and minimal amongst the p-adic valuations of all coefficients of $Q_1(x)$. But now, if $p^r > \deg(Q_0(x)), \deg(Q_1(x))$, it is easy to check that the p-adic valuation of $[1, 2, \ldots, n]a_n$ is negative for $n = p^r + m$ and a_n the coefficient of x^n in $f(x)$. △

Remark 2.7.6. The resort to the Hermite–Lindemann–Weierstrass theorem and Mahler's (partial) p-adic analog is not accidental in the proof of Theorem 2.7.2. In fact, reversing the logic at least in part, the statement of the theorem implies, for example, the irrationality of $\log(1-1/n)$ for all integers $n \in \mathbf{Z} \setminus \{1\}$; for if this (archimedean) logarithm took a rational value p/q, then

$$f(x) := \frac{q\log(1-x) - p}{1-nx} = q\frac{\log(1-x) - \log(1-1/n)}{1-nx} \in \mathbf{Q}[\![x]\!] \cap \mathcal{O}\left(\mathbf{C} \setminus [1,\infty)\right)$$

would meet the integrality and holomorphy constraints in Theorem 2.7.2, but the rational functions $Q_0(x), Q_1(x) \in \mathbf{Q}(x)$ in the expression $f(x) = Q_0(x)+Q_1(x)\log(1-x)$ would be singular at $x = 1/n$, and thus definitely not from the ring $\mathbf{Q}\left[x, \frac{1}{1-x}\right]$. We shall return to this type of issue in § 15.2. △

2.7.7. *Theorem 2.7.1 as a refinement of the Borel–Pólya–Zudilin rationality criterion.* We make three remarks about Theorem 2.7.1. First, in the discussion in Basic Remark 2.1.2 we do indeed recover the more precise rationality statement in the original Borel–Pólya theorem, for we can have $\tau(\mathbf{b}) = \tau(\mathbf{0}) = 0$ in that setting. Second, on a given simply connected domain $\Omega \ni 0$ of the complex plane with conformal mapping radius $\rho(\Omega, 0) > 1$, all transcendental $\mathbf{Q}[\![x]\!]$ formal function germs with a denominator type of the form

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\ldots,b_1 n]\cdots[1,\ldots,b_r n]}, \qquad a_n \in \mathbf{Z}, \tag{2.7.8}$$

must meet a denominator type gap

$$b_1 + \ldots + b_r \geq (2/3)\log\rho(\Omega, 0). \tag{2.7.9}$$

If there are at least $m \geq 2$ such $\mathbf{Q}(x)$-linearly independent functions, the coefficient $2/3$ in (2.7.9) improves to $m/(m+1)$.

Finally, in the most basic situation of all taking Ω to be a round disc $|x| < R$ of a radius $R > e^{\sigma}$ centered at 0 (as considered by Apéry, except that now — like Borel — we assume meromorphy rather than holomorphy), we explain how Theorem 2.7.1 implies that $f(x) \in \mathbf{Q}(x)$. Applying Theorem 2.7.1 directly, we deduce to start with that the corresponding $\mathbf{Q}(x)$-vector space $\mathcal{H}$ generated by such functions is finite dimensional, and in particular consists of holonomic functions. However, if $f(x) = \sum a_n x^n \in \mathbf{Q}[\![x]\!]$ is meromorphic on Ω, then, with $\zeta = e^{2\pi i/m}$, so are the twists

$$\frac{1}{m}\sum_{i=0}^{m-1} f(\zeta^i x)\zeta^{-ik} = \sum a_{mn+k}x^{nm+k},$$

and those have the same denominator type as $f(x)$. It follows that $\mathcal{H}$ is preserved by $x \mapsto \zeta x$ for any m. The (non-apparent) singularities of the corresponding differential equation cannot be invariant under all these rational rotations unless they are a subset of $\{0, \infty\}$. But this implies that any such $f(x)$ must be meromorphic on $\mathbf{C}$, and (after clearing denominators) we may apply Theorem 2.7.1 again, taking now R to be arbitrarily large, to deduce that $\dim_{\mathbf{Q}(x)} \mathcal{H} = 1$. (Note that there do exist finite-dimensional $\mathbf{Q}(x)$-vector spaces of dimension greater than 1 which are generated by $\mathbf{Z}[\![x]\!]$ holomorphic functions on $\mathbf{D}$ and are invariant under $x \mapsto \zeta x$ for all rational rotations; for example, the $\mathbf{Q}(x)$-vector space generated by 1 and $f(x) = \sum x^{n!}$. The latter, of course, is non-holonomic.)

We can summarize the three remarks by the following refinement of the Borel–Pólya rationality criterion, and also of Zudilin's determinantal criterion [Zud17].

Theorem 2.7.10. *Consider a contractible open domain $\Omega \ni 0$ in the complex plane and a formal power series of the arithmetic type*

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\ldots,b_1 n]\cdots[1,\ldots,b_r n]}, \qquad a_n \in \mathbf{Z}, \quad \forall n \in \mathbf{N}, \tag{2.7.11}$$

which is the $x = 0$ germ of a meromorphic function on Ω. Suppose that either

(i) Ω is a round disc $|x| < R$ of a radius $R > \exp(b_1 + \ldots + b_r)$; or else that
(ii) the conformal mapping radius $\rho(\Omega, 0)$ of Ω at the origin exceeds

$$\exp\left(\frac{3}{2}(b_1 + \ldots + b_r)\right).$$

Then $f(x) \in \mathbf{Q}(x)$ is the Taylor expansion of a rational function. □

2.8. **Arithmetic characterizations beyond the logarithm.** In light of Theorem 2.7.2, it is natural to inquire of arithmetic characterizations of other basic transcendental functions in terms of their domains of analyticity and the arithmetic behavior of their power series. In view of Belyĭ's theorem [BG06, § 12.3], a natural place to start is (as in Theorem 2.7.2) with power series that can be analytically continued as *multivalued* holomorphic functions along all paths in $\mathbf{P}^1 \setminus \{0, 1, \infty\}$. Going further than the denominator type $[1, \ldots, n]$ of Theorem 2.7.2 requires to use a multivalent map φ, but there is still a local univalence input, discussed in § 2.9 and formalized in § 9.0.12 and Corollary 9.0.19, which is essential for our approach to irrationality proofs.

In any case, if $\tau = \tau(\mathbf{b})$, a necessary condition for our methods to have any hope of applying is that $|\varphi'(0)| > e^{\tau}$. A theorem of Carathéodory [Car54, (412.8) on page 198] shows that $|\varphi'(0)| \leq 16$ for all holomorphic maps $\varphi : \mathbf{D} \to \mathbf{C} \smallsetminus \{1\}$ subject to $\varphi^{-1}(0) = \{0\}$, with equality holding if and only if $\varphi(z) = \lambda(cz)$ with $|c| = 1$. Hence, in this setting, it is necessary that $\lambda'(0) = 16 > e^{\tau}$ (see § 2.9 for details on why the specific assumptions in Carathéodory's theorem is relevant). This necessary condition is certainly met by $\tau = 2$. In particular, Corollary 2.6.1 implies that the $\mathbf{Q}(x)$-vector space of type $[1, \ldots, n]^2$ functions holonomic on $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ is finite dimensional. As we shall explain below, our method of proof for both Theorems A and C can be summarized as making a sufficient way towards the determination of that finite-dimensional space.

Conjecture 2.8.1. *The following conditions on a formal power series $f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbf{Q}[\![x]\!]$ convergent in $|x| < 1$ are equivalent:*

1. *$f(x)$ is analytically continuable as a holomorphic function to $\mathbf{C} \smallsetminus [1, \infty)$, and furthermore* as a meromorphic function *along all paths in $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$; and there is an $M \in \mathbf{N}_{>0}$ such that $[1, \ldots, n]^2 a_n \in M^{-1}\mathbf{Z}$ for all $n \in \mathbf{N}$.*
2. *There are rational functions $Q_0, \ldots, Q_4 \in \mathbf{Q}\left[x, \dfrac{1}{1-x}\right] \subset \mathbf{Q}(x)$ with*

$$f(x) = Q_0(x) + Q_1(x)\log(1-x) + Q_2(x)\log^2(1-x) + Q_3(x)\mathrm{Li}_2(x) + \frac{Q_4(x)}{\sqrt{1-x}} \int_0^x \frac{\log(1-t)}{t\sqrt{1-t}}\, dt.$$

Here, $\mathrm{Li}_2(x) := -\int_0^x \log(1-t)\, d\log t = \sum_{n=1}^{\infty} x^n/n^2$ is the standard dilogarithm function branch, and the hypothetical solution space is discussed in more depth in § 10.1. One should compare this conjecture to Theorem 2.7.2. In either case, one may consider a bipartite approach. The first part is to devise a setup in Theorem 2.5.1 that proves the finite-dimensionality of the $\mathbf{Q}(x)$-vector space of such functions; the second part is to give a bound for this space which coincides with the number of known functions. The fact that $16 > e$ (respectively $16 > e^2$) establishes the first claim in either case. In the second case, however, the best bound on the dimension we can currently establish is 9 rather than 5 (see Remark A.5.2 and Equation A.5.3). Ruling any possible further functions out remains a difficult problem currently beyond the reach of our methods in this paper.

Remark 2.8.2. Similarly to Remark 2.7.6, the $\mathbf{Q}[x, 1/(1-x)]$ refinement contains, like a hidden particular clause in this form of Conjecture 2.8.1, the $\mathbf{Q}$-linear independence of the $x = 1/n$ special values of the five functions 1, $\log(1-x)$, $\log^2(1-x)$, $\mathrm{Li}_2(x)$, and $\frac{1}{\sqrt{1-x}} \int_0^x \frac{\log(1-t)}{t\sqrt{1-t}}\, dt$ in the statement of the conjecture, for every $n \in \mathbf{Z} \smallsetminus \{0, 1\}$. If for example $\mathrm{Li}_2(1/n) \in \mathbf{Q}$, then the point is that the function

$$f(x) := \frac{\mathrm{Li}_2(x) - \mathrm{Li}_2(1/n)}{1 - nx} \in \mathcal{O}(\mathbf{C} \smallsetminus [1, \infty)) \cap \mathbf{Q}[\![x]\!]$$

meets all the conditions in (1) of the conjecture, but it is manifestly not contained in the solution $\mathbf{Q}[x, 1/(1-x)]$-module prescribed by (2).

For all $|n| \geq N_0$, where N_0 is some (large, explicitly computable) number, the $\mathbf{Q}$-linear independence of the $x = 1/n$ special values of those five functions follows as a very particular case of the general Theorem 15.1.3 from the theory of special values of G-functions. For the dilogarithm function, the first such result was proved already by Maier [Mai27, § 8], in a work that foreshadowed (and directly inspired) Siegel's 1929 paper [Zan14]. The irrationality $\mathrm{Li}_2(1/n) \notin \mathbf{Q}$ has at present only been proved [Hat93, RV05, RV19] for $n \notin \{-4,-3,-2,2,3,4,5\}$. The issue is discussed further in § 15.2. △

In the main spirit of our paper, one could even ask for variations of Conjecture 2.8.1 that allow for further (possible) singularities in the convergence disc $|x| < 1$ of the original branch $f(x)$, for example:

Question 2.8.3. *Do the conclusions of Conjecture 2.8.1 still hold if the meromorphic continuability on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ in condition (1) is relaxed to a meromorphic continuability on $\mathbf{P}^1 \smallsetminus \{0,\delta,1,\infty\}$, for some $\delta \in [-1/2,1/2]$?*

In particular, for a power series $f(x) \in \mathbf{Q}[\![x]\!]$ convergent on $|x| < 1$ and defining holonomic functions on $\mathbf{P}^1 \smallsetminus \{0,\delta,1,\infty\}$ of the denominators type condition $a_n[1,\ldots,n]^2 \in \mathbf{Z}$, where $\delta \in [-1/2,1/2]$ is an arbitrary fourth puncture, does $f(x)$ automatically extend through that fourth puncture $x = \delta$ to define a holonomic function on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$?

While these questions seem rather awkward for our method of rational holonomy bounds as developed in this paper, we are able to fully resolves a sub-problem intermediate in difficulty between Theorem 2.7.2 and Conjecture 2.8.1, namely when the denominator type has the form $[1,2,\ldots,n][1,2,\ldots,n/2]$, that is "a case of $\tau = 3/2$" where the first new function after $\log(1-x)$ pops out, namely, the function $\log^2(1-x)$. Here Conjecture 2.8.1 becomes the $\delta = 0$ case of the following theorem, responding affirmatively to Question 2.8.3 for the subcase of $[1,\ldots,n][1,\ldots,n/2]$ types:

Theorem 2.8.4. *Suppose $f(x) = \sum_{n=0}^{\infty} a_n x^n \in \mathbf{Q}[\![x]\!]$ has $[1,\ldots,n][1,\ldots,n/2]a_n \in \mathbf{Z}$ for all $n \in \mathbf{N}$, is holomorphic in $\mathbf{C} \smallsetminus [1,\infty)$, and is analytically continuable as a meromorphic function along all paths in $\mathbf{P}^1 \smallsetminus \{0,\delta,1,\infty\}$, for some $\delta \in (-\infty,1)$.*

Then

$$f(x) = Q_0(x) + Q_1(x)\log(1-x) + Q_2(x)\log^2(1-x) \tag{2.8.5}$$

for some rational functions of the form $Q_0, Q_1, Q_2 \in \mathbf{Q}\left[x, \frac{1}{1-x}\right] \subset \mathbf{Q}(x)$.

In particular,

$(*)$ *$f(x)$ continues analytically as a meromorphic function along all paths in $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$.*

Some immediate applications of part $(*)$ of Theorem 2.8.4 to $\mathbf{Q}$-linear independence proofs are treated in § 2.11.12 further down in this introduction, as a proof-of-concept for our method. It is there that $(*)$ is proved, as an application of Theorem 2.5.1. To conclude the full Theorem 2.8.4 requires a subtler holonomy bound and it is carried out in § 6.8.

2.9. **Overconvergence and univalent leaves.** We now turn to the basic mechanism for irrationality proofs by extending the method of Apéry limits. We will follow this in § 2.10 with some explicit examples, and in § 2.11 with a proof-of-concept application to some new $\mathbf{Q}$-linear independence proofs.

Consider $\Sigma \subset \mathbf{D}_R := \{x \in \mathbf{C} : |x| < R\}$ a discrete subset of the open complex disc of radius $R \in (0,\infty]$ (possibly including the disc center $x = 0$), and $f(x) \in \mathbf{C}[\![x]\!]$ a holomorphic function germ at the center point that continues analytically as a holomorphic function along all paths in $\mathbf{D}_R \smallsetminus \Sigma$. Let us define the subset $\Sigma_f^+ \subset \Sigma$, to necessarily include 0 if $0 \in \Sigma$, to consist of those $\beta \in \Sigma$ for which the radial analytic continuation of $f(x) \in \mathbf{C}[\![x]\!]$ from $x = 0$ towards $x = \beta$ remains bounded. We say that the power series $f(x)$ is *overconvergent* at Σ_f^+ and *extends to* $\mathbf{D}_R \smallsetminus \Sigma$ *as a multivalued holomorphic function*. Then the radius of convergence of the initial power series germ $f(x) \in \mathbf{C}[\![x]\!]$ is equal to $\min_{\beta \in (\{R\} \cup \Sigma) \smallsetminus \Sigma_f^+} |\beta|$. The following trivial lemma is crucial for our approach to Theorems A and C; we note that this type of statement on compatibility with integrations becomes completely false if we replace *holomorphic* by *meromorphic* everywhere in the previous paragraph.

Lemma 2.9.1. *There is an equality* $\Sigma^+_{\int_0^x f(t)\,dt} = \Sigma_f^+$.

Given now a holomorphic mapping $\varphi : \mathbf{D} \to \mathbf{D}_R$ with $\varphi(0) = 0$, we can apply the same notion to the pulled-back power series $f(\varphi(z)) \in \mathbf{C}[\![z]\!]$, which is a $z = 0$ holomorphic function germ that extends to $\mathbf{D} \smallsetminus \varphi^{-1}(\Sigma)$ as a multivalued holomorphic function. In general, there is no relationship between the overconvergence sets $\Sigma_f^+ \subset \Sigma \subset \mathbf{D}_R$ and $\big(\varphi^{-1}(\Sigma)\big)^+_{\varphi^* f} \subset \varphi^{-1}(\Sigma) \subset \mathbf{D}$ for f and $\varphi^* f$.

But suppose there is a contractible open neighborhood $0 \in \Omega \in \mathbf{D}$ on which $\varphi|_\Omega : \Omega \xrightarrow{\simeq} \varphi(\Omega)$ is univalent and, therefore, a conformal isomorphism onto the image open neighborhood $\varphi(U) \ni 0$. Assume furthermore that $\varphi^{-1}(0) = \{0\}$ and that each point in $\varphi(\Omega) \cap \Sigma_f^+$ has exactly one pre-image under the analytic map φ, that is:

$$\varphi^{-1}\left(\varphi(\Omega) \cap \Sigma_f^+\right) \subset \Omega.$$

Then, in particular, $f(\varphi(z))$ is holomorphic on at least Ω:

$$\left(\varphi^{-1}\left(\Sigma\right) \cap \Omega\right)^+_{\varphi^* f} = \varphi^{-1}\left(\Sigma_f^+\right) \cap \Omega = \varphi^{-1}\left(\Sigma_f^+ \cap \varphi(\Omega)\right). \tag{2.9.2}$$

This is the univalence input we alluded to. If now, in addition, $\varphi(\mathbf{D}) \cap \Sigma = \Sigma_f^+ \cap \varphi(\Omega)$, it follows at once that the multivalued holomorphic function $f(\varphi(z))$ on $\mathbf{D} \smallsetminus \varphi^{-1}(\Sigma) = \mathbf{D} \smallsetminus \varphi^{-1}\left(\Sigma_f^+ \cap \varphi(\Omega)\right) = \mathbf{D} \smallsetminus \left(\varphi^{-1}\left(\Sigma\right) \cap \Omega\right)^+_{\varphi^* f}$ is in fact a (single-valued) holomorphic function on the whole disc $\mathbf{D}$, that is a convergent power series on that disc.

We summarize the basic property that we just proved:

Proposition 2.9.3. *Let $f \in \mathbf{C}[\![x]\!]$ be a holomorphic function germ which extends as a multivalued holomorphic function on the Riemann surface $\mathbf{P}^1 \smallsetminus \Sigma$, for some finite set of punctures Σ on the Riemann sphere. Consider a disjoint partition $\Sigma = \Sigma^0 \sqcup \Sigma^1$, a holomorphic map $\varphi : \mathbf{D} \to \mathbf{P}^1 \smallsetminus \Sigma^1$ that takes $\varphi(0) = 0$, and a contractible open neighborhood $0 \in \Omega \subset \mathbf{D}$ on which φ restricts as a univalent map (equivalently: $\varphi_{|\Omega} : \Omega \xrightarrow{\simeq} \varphi(\Omega)$ is a conformal isomorphism). We assume that $\varphi^{-1}\left(\Sigma^0\right) \subset \Omega$ and that $f \in \mathcal{O}(\varphi(\Omega))$ is holomorphic on $\varphi(\Omega)$.*

Then, the pulled-back germ $f(\varphi(z)) \in \mathbf{C}[\![z]\!]$ converges on the full disc $\mathbf{D}$.

Remark 2.9.4. The assumptions on the triple $(\varphi, \Omega, \Sigma)$ in Proposition 2.9.3 can alternatively, and slightly more succinctly, be summarized by having a holomorphic mapping $\varphi : (\mathbf{D}, 0) \to (\mathbf{C}, 0)$ that restricts univalently on the contractible open neighborhood $\Omega \ni 0$, and such that $\varphi^{-1}(\Sigma) \subset \Omega$ for the finite puncture set Σ. We chose the formulation with $\Sigma = \Sigma^0 \sqcup \Sigma^1$ to highlight the practical presence of universal maps φ when the singularity type (Σ^0, Σ^1) is given but the open neighborhood $\Omega \ni 0$ is kept unspecified.

2.9.5. *The modular lambda map.* In this general context, the significance of the modular lambda map (1.2.8) is in the observation that $\varphi(z) := \lambda(z)$ is the universal map in Proposition 2.9.3 for the case $\Sigma^0 = \{0\}$ and $\Sigma^1 = \{1, \infty\}$ (upon keeping fluid the choice of an unspecified open neighborhood $\Omega \ni 0$). Its derivative $\lambda'(0) = 16$ therefore maximizes the conformal size of any such map. This can be considered (see Remark 2.11.3 for a direct connection) as the multivalent analog of the role of the domain $\Omega = \mathbf{C} \smallsetminus [1, \infty)$ and the Koebe map $\varphi(z) = 4z/(1+z)^2$ in the proof of Theorem 2.7.2. Concretely, if $f(x) \in \mathbf{C}[\![x]\!]$ continues analytically along all paths as a holomorphic function on $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ (an example is any balanced hypergeometric series), then $f(\lambda(z)) \in \mathbf{C}[\![z]\!]$ converges on the open unit disc $z \in \mathbf{D}$. A basic illustration is the classic Jacobi formula

$$\sum_{n=0}^{\infty} \binom{2n}{n}^2 \left(\frac{\lambda(q)}{16}\right)^n = \left(\sum_{n=0}^{\infty} q^{n^2}\right)^2,$$

where the holonomicity in $x = \lambda(q)$ is an expression of the Picard–Fuchs ODE for the de Rham cohomology of the Legendre elliptic curve over

$$Y(2)_{\mathbf{C}} = \operatorname{Spec} \mathbf{C}\left[x, 1/x, 1/(1-x)\right].$$

Our proof of Theorem A will involve a similar expression § 11.1 of the Picard–Fuchs ODE over the modular curve

$$Y_0(6)_{\mathbf{C}} \cong \operatorname{Spec} \mathbf{C}\left[x, 1/x, 1/(1-x), 1/(1-9x)\right],$$

in which $\zeta(2)$ and $L(2,\chi_{-3})$ emerge as the Eichler periods.

2.10. **First irrationality proofs.** In Remark 2.8.2 on Conjecture 2.8.1, we observed that the prescribed $\mathbf{Q}[x, 1/(1-x)]$-module has direct irrationality implications on special values at points of the form $x = 1/n$. However, the method of our present paper only addresses $\mathbf{Q}(x)$-vector spaces in the framework of Theorem 2.5.1, but not their integral structures over finitely generated $\mathbf{Q}$-algebras intermediate between $\mathbf{Q}[x]$ and $\mathbf{Q}(x)$. We now explain how even the cruder $\mathbf{Q}(x)$-form of Conjecture 2.8.1 (as enhanced by Question 2.8.3) casts a method for establishing irrationality proofs. These are now in the form of Apéry limits, as opposed to the straight special values of the functions in the relevant holonomic module.

The following expands upon what we have already discussed in the introduction. The ideal situation is as follows. Given an interesting period η, one writes down a holonomic function $f(x)$ with coefficients in $\mathbf{Q}(\eta)$. Assuming for the contradiction that $\eta \in \mathbf{Q}$ and hence $f(x) \in \mathbf{Q}[\![x]\!]$, this function (together with its derivatives) provides a space of holonomic functions of some explicit denominator type and dimension over $\mathbf{Q}(x)$. In addition, depending on the circumstances, there will also exist other known functions in this space. Considerations of monodromy (or otherwise) typically allow one to show that this space of known functions is $\mathbf{Q}(x)$ (and even $\mathbf{C}(x)$)-linearly independent from the functions coming from $f(x)$. If the lower bound coming from the span of such functions exceeds the upper bounds from our theorem, we obtain the desired irrationality of η.

Consider, for instance, our task to establish the $\mathbf{Q}$-linear dependence $1, \zeta(2)$, and $L(2,\chi_{-3})$. The ideal scenario would be to use a putative $\mathbf{Q}$-linear dependence to write down such an $f(x)$ with denominators of type $\tau = [1, 2, \ldots, n]^2$ which extends holomorphically along all paths in $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$, but such that $f(x)$ is *not* in the $\mathbf{Q}(x)$-vector space generated by the five functions in Conjecture 2.8.1. Then

Conjecture 2.8.1 would immediately give a contradiction. This is not possible, but clearly we can get away with something weaker. As mentioned, we *can* prove a bound of 9 on the dimension of such functions. Now such a bound would still be sufficient as long as the span of $f(x)$ and its derivatives were linearly independent from these five functions and gave a complementary $\mathbf{Q}(x)$-vector space of dimension at least 5. In practice, even this fails in two respects. First, the function $f(x)$ we construct only generates a holonomic module of dimension 4. Second, the function $f(x)$ has additional singularities at paths in $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ to both $\delta = 1/9$ and $\delta = -1/8$. It turns out that we can still bound the space of functions by 9 with these additional singularities, but the numerology still falls *just* short of our desired application. Instead, we have to additionally also include *integrations* of these (and other) functions into our story, and this is how we ultimately achieve the proof of Theorem A, which is perhaps the most subtle of our applications. It seems useful, however, to give examples where the approach as described above works directly, first by reproving the (known by Lambert in 1761!) irrationality of $\log 3$, and then (in Theorem 2.11.17) to devising a new irrationality result.

Basic Remark 2.10.1. Turning now to the main style of applications of holonomy bounds to irrationality proofs, the following is a simple example due to Zudilin [Zud17, § 3], in which case (ii), but not case (i) of Theorem 2.7.10 provides an irrationality proof of the period $\log 3$ out of the consideration of the integrals

$$\begin{aligned} f(x) &:= \frac{1}{\sqrt{1-4x+x^2}} \int_{2-\sqrt{3}}^{x} \frac{dt}{\sqrt{1-4t+t^2}} \\ &= \frac{1}{2} \sum_{n=0}^{\infty} (b_n - a_n \log 3) \cdot \frac{x^n}{2^n}, \qquad a_n \in \mathbf{Z}, \quad [1, \dots, n]\, b_n \in \mathbf{Z}, \quad \forall n \in \mathbf{N}. \end{aligned} \tag{2.10.2}$$

For the second line, we use the binomial expansion of $(1-4x+x^2)^{-1/2} \in \mathbf{Z}[\![x/2]\!]$ and the fact that $-\log\left(2 - x + \sqrt{1-4x+x^2}\right)$ is an explicit primitive of $1/\sqrt{1-4x+x^2}$.

Of course, $f(x)$ is not a G-function as it has transcendental coefficients from involving $\log 3$; rather, it is a $\mathbf{C}$-linear combination of two G-functions on $\mathbf{P}^1 \smallsetminus \{2 \pm \sqrt{3}, \infty\}$, and $\log 3 \in \mathbf{C}$ gets characterized as the unique (holonomic) coefficient in such a combination to give a branch regular (holomorphic) at the smaller singularity $x = 2 - \sqrt{3}$. (This is rather transparently revealed by the fact that both factors in (2.10.2) switch sign after a simple loop going around that singularity $x = 2 - \sqrt{3}$, and thus their product has no monodromy at $x = 2 - \sqrt{3}$.)

But we can turn this around and get an irrationality proof of $\log 3 \notin \mathbf{Q}$ as an application of Theorem 2.7.10 (and, hence, ultimately of the univalent holonomy bound). Proving $\log 3 \notin \mathbf{Q}$ means precisely proving that $f(2x) \notin \mathbf{Q}[\![x]\!]$. Suppose not. Then $f(2x) \in \mathbf{Q}[\![x]\!]$ has, upon clearing a fixed positive integer denominator, visibly the type (2.7.11) with $r = 1$ and $(b_1, \dots, b_r) = (1)$. At the same time, by construction we have $f(2x)$ holomorphic on the domain

$$x \in \Omega := \mathbf{C} \smallsetminus \left[\frac{2+\sqrt{3}}{2}, \infty\right)$$

of Riemann mapping radius

$$\rho(\Omega, 0) = 2(2+\sqrt{3}) = 7.4641\ldots > 4.481689\ldots = e^{3/2}.$$

Theorem 2.7.10 (ii) proves that every such function $f(2x)$ has to be a rational function. Obviously, the putative function from (2.10.2) (which would only have existed had $\log 3$ been a rational number) is not rational, and so contrapositively this argument gives a proof of the irrationality of $\log 3$.

And yet, as $2+\sqrt{3} = 3.73205\ldots < 5.43656\ldots = 2e$, the $2^n[1,\ldots,n]$ denominators growth rate in these approximations $\log 3 \approx b_n/a_n$ exceeds the reciprocal of the decay rate $2-\sqrt{3}$ of the error $|\log 3 - b_n/a_n|$ of the approximations. In other words, case (ii) applies in the theorem, whereas case (i) does not. Thus we find an irrationality proof, by G-function methods, without actually constructing any rapidly convergent explicit (*holonomic*) rational approximants. (A more complicated construction [Sal07, Sor16] to pass the latter requirement is known in the case of $\log 3$, but not for say $\log p$ where p is any sufficiently big prime.)

As we will see, the usefulness of Theorem 2.5.1 lies in the possibility of using — instead of domains $\Omega \subset \mathbf{C}$ as on this example — *multivalent* mappings such as $\varphi(z) := \lambda(z)$, a holomorphic function on $\mathbf{D}$ whose derivative $|\varphi'(0)| = 16 > e^2$ fortuitously exceeds the growth rate of the $[1,\ldots,n]^2$ layer of denominators common to several linear independence problems of interest here (including the case of Theorems A and C), and which applies to the holonomic functions on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$. △

2.11. **The multivalent case: first new linear independence results.** In this section, we prove half of Theorem 2.8.4 — namely, part $(*)$ — using a multivalent map φ in Theorem 2.5.1, and derive as a consequence a first $\mathbf{Q}$-linear independence proof which, unlike with Basic Remark 2.10.1 to which it is otherwise entirely similar, is actually a new result. The second half of Theorem 2.8.4, which is irrelevant to this application, will be proved in § 6.8.

A key point to observe is that we shall definitely need a multivalent choice for φ.

Basic Remark 2.11.1. Koebe's quarter theorem states that $|\varphi'(0)| \leq 4$ for all univalent holomorphic maps of pointed domains $\varphi : (\mathbf{D}, 0) \to (\mathbf{C} \smallsetminus \{1\}, 0)$, and that equality holds if and only if $\varphi(z) = G(cz)$ with $|c| = 1$, where

$$G(z) := \frac{4z}{(1+z)^2} = 1 - \left(\frac{1-z}{1+z}\right)^2$$

is *Koebe's extremal function*, the Riemann uniformization map at the origin of the slit complex plane $\mathbf{C} \smallsetminus [1,\infty)$. In particular, if we restrict φ to univalent maps in Theorem 2.5.1 then we cannot hope to prove Theorem 2.8.4 since then

$$|\varphi'(0)| \leq 4 < 4.481689\ldots = e^{3/2}.$$

The Koebe map is $1:1$ on the open unit disc but it extends to a $2:1$ rational map $\mathbf{C} \smallsetminus \{\pm 1\} \to \mathbf{C} \smallsetminus \{1\}$. Pre-composing this quadratic rational map with the Riemann uniformization map $\mathbf{D} \to \mathbf{C} \smallsetminus ((-\infty,-1] \cup [1,\infty))$, which is simply the map $\sqrt{G(z^2)} = 2z/(1+z^2)$, we end up with the bivalent map

$$\varphi : \mathbf{D} \to \mathbf{C} \smallsetminus \{1\}, \qquad \varphi(z) := G\left(\sqrt{G(z^2)}\right) = \frac{8(z+z^3)}{(1+z)^4} = 1 - \left(\frac{1-z}{1+z}\right)^4. \tag{2.11.2}$$

In the present section, analogously to the role of Koebe's univalent map for the proof of Theorem 2.7.2, we make a use of the basic properties of the bivalent map (2.11.2). This map bijects $(-1,1) \xrightarrow{\simeq} (-\infty,1)$ and is bivalent on $\mathbf{D} \smallsetminus (-1,1)$, taking either of the two connected halves conformally isomorphically onto $\mathbf{C} \smallsetminus (-\infty, 1]$. This shows in particular that the case $\varphi : (\mathbf{D}, 0) \to (\mathbf{C} \smallsetminus \{1\}, 0)$ in Proposition 2.9.3 with $\Sigma^1 = \{1, \infty\}$ and an *arbitrary* $\Sigma^0 \subset (-\infty, 1)$ can have a derivative as big as $|\varphi'(0)| = 8$. △

The continuation of this construction explains the central role of the modular lambda map in our paper:

Remark 2.11.3. We can repeat the process of getting from $G(z) = 4z/(1+z)^2$ to $G\left(\sqrt{G(z^2)}\right) = 8(z+z^3)/(1+z)^4$ by post-composing next with the Riemann map of the complement in $\mathbf{C}$ of the union of the four normal external rays out from the fourth roots of unity $z = \pm i$ (the points that give additional zeros of the map (2.11.2), that we want to avoid having for $\Sigma^0 = \{0\}$) and $z = \pm 1$ (which give values 1 and ∞ for (2.11.2), which we want to avoid having for $\Sigma^1 = \{1, \infty\}$). But the Riemann map of this $\mathbf{Z}/4$-rotationally symmetrically slit region is just $\sqrt[4]{G(z^4)}$. The result is the quadrivalent map

$$\varphi : \mathbf{D} \to \mathbf{C} \smallsetminus \{1\}, \qquad \varphi(z) := G\left(\sqrt{G\left(\sqrt{G(z^4)}\right)}\right) = \frac{8\sqrt{2}\,(1+z^2)^2\sqrt{1+z^4}}{\left(z\sqrt{2} + \sqrt{1+z^4}\right)^4}, \tag{2.11.4}$$

which has the bigger derivative $\varphi'(0) = 8\sqrt{2}$ while still serving in Proposition 2.9.3 for the case $\Sigma^0 = \{0\}$ and $\Sigma^1 = \{1, \infty\}$.

Continuing these iterations, we find that the nesting with n square roots

$$\varphi_n : \mathbf{D} \to \mathbf{C} \smallsetminus \{1\}, \qquad \varphi_n(z) := G\left(\sqrt{G\left(\sqrt{G\left(\sqrt{\cdots G(z^{2^n})}\right)}\right)}\right) \tag{2.11.5}$$

continues to serve in the $\Sigma^0 = \{0\}, \Sigma^1 = \{1, \infty\}$ case of (2.9.3), while having the derivative

$$\varphi_n'(0) = 4^{1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}} = 16^{1-2^{-n-1}}. \tag{2.11.6}$$

This constructs a sequence of 2-solvable algebraic power series $\varphi_0(q), \varphi_1(q), \varphi_2(q), \ldots$ in $\mathbf{C}[\![q]\!]$ starting with the Koebe map $\varphi_0(q) = 4q/(1+q)^2$ and converging coefficients-wise, as well as locally uniformly on $q \in \mathbf{D}$, to the modular lambda map (1.2.8). The latter fact was known in essence to Landen, Legendre, and Gauss [BB98, § 1] in the form of the *arithmetic-geometric mean iteration* $(a, b) \rightsquigarrow \left(\frac{a+b}{2}, \sqrt{ab}\right)$. △

We base our proof of Theorem 2.8.4 on the bivalent example

$$\varphi(z) := G\left(\sqrt{G(z^2)}\right) = 8(z+z^3)/(1+z)^4$$

from Basic Remark 2.11.1. Crucially, the restriction to $\mathbf{D}$ of this rational map has the fairly big derivative $\varphi'(0) = 8$ all the while inducing a bijection $(-1, 1) \xrightarrow{\simeq} (-\infty, 1)$ and conformal isomorphisms $\mathbf{D} \cap \{\mathrm{im}(z) > 0\} \xrightarrow{\simeq} \mathbf{C} \smallsetminus (-\infty, 1]$ and $\mathbf{D} \cap \{\mathrm{im}(z) < 0\} \xrightarrow{\simeq} \mathbf{C} \smallsetminus (-\infty, 1]$.

The rationality of this basic function also allows for an explicit formula of the double integral occurring in the holonomy bound (2.5.4).

Lemma 2.11.7. *The Bost–Charles double integral of the map*

$$\varphi(z) := 8(z+z^3)/(1+z)^4$$

has the following explicit evaluation:

$$\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) = \log 8 + \frac{4G}{\pi}, \tag{2.11.8}$$

where $G := L(2,\chi_{-4})$ is the Catalan constant.

Proof. This time, to compare with the proof of the univalent case of Theorem 2.7.1, we have the factorization

$$\frac{\varphi(z)-\varphi(w)}{z-w} = 8\frac{(1-zw)(1+ix-iy-xy)(1-ix+iy-xy)}{(1+z)^4(1+w)^4}, \tag{2.11.9}$$

which does have zeros on the unit polydisc $\mathbf{D}^2$: the Bost–Charles *overflow* [BC22, § 5] is positive, and it equals the Mahler measure

$$\begin{aligned} m\left(1+x^2+y^2-4xy+x^2y^2\right) &= m(1+ix-iy-xy)+m(1-ix+iy-xy)\\ &= 2m(1+x+y-xy) = 4G/\pi, \end{aligned}$$

where the two integrals make the respective unimodular change of variables $(x,y) \rightsquigarrow (\pm ix, \mp iy)$, and the last evaluation is due to Smyth [Boy98]. □

We will divide the proof of Theorem 2.8.4 into two parts: property $(*)$ in our statement of the theorem, regarding the meromorphic extendability through δ in all analytic continuations; and the derivation of the full form (2.8.5) granting $(*)$. We now prove the first part — property $(*)$ — and derive from it a showcase application in § 2.11.12 to $\mathbf{Q}$-linear independence. The second part is subtler and will be proved in § 6.8 based on the refined holonomy bound Theorem 6.0.2.

Proof of part $(*)$ in Theorem 2.8.4. Suppose to the contrary that there exists a $\delta \in (-\infty,1) \smallsetminus \{0\}$ such that $f(x)$ converges on $|x|<1$ while having analytic continuations along all paths in $\mathbf{P}^1 \smallsetminus \{0,\delta,1,\infty\}$ and with eventually a nontrivial local monodromy around $x=\delta$. Consider the Möbius involution $x \mapsto x/(x-1)$ that fixes the origin of the expansions, preserves the $[1,\dots,n][1,\dots,n/2]$ denominators type, exchanges the punctures $1 \leftrightarrow \infty$, and maps $\delta \leftrightarrow \delta/(\delta-1)$ to a different puncture (since $\delta \neq 2$), which is also in $(-\infty,1)$. Then the formal power series $f\big(x/(x-1)\big) \in \mathbf{Q}[\![x]\!]$ has similar properties to $f(x)$, except now for having meromorphic continuations along all paths in $\mathbf{P}^1 \smallsetminus \big\{0,\delta/(\delta-1),1,\infty\big\}$ and with eventually a nontrivial local monodromy around $x=\delta/(\delta-1)$. As $\delta \neq \delta/(\delta-1)$, it follows at once that the following five functions, all of the $[1,\dots,n][1,\dots,n/2]$ denominator type, are $\mathbf{C}(x)$-linearly independent:

$$1, \quad \log(1-x), \quad \log^2(1-x), \quad f(x), \quad f\left(\frac{x}{x-1}\right). \tag{2.11.10}$$

We use Theorem 2.5.1 with the 5×2 array

$$\mathbf{b} := \begin{pmatrix} 0 & 1 & 1 & 1 & 1\\ 0 & 0 & \frac12 & \frac12 & \frac12 \end{pmatrix}^{\mathrm{t}},$$

corresponding to the denominator types in the ordered list (2.11.10). We calculate

$$\tau(\mathbf{b}) = \frac{1\cdot 0 + 3\cdot 1 + (5+7+9)\cdot(3/2)}{5^2} = \frac{69}{50} = 1.38. \tag{2.11.11}$$

For the map φ, we select

$$\varphi(z) := \frac{8(z+z^3)}{(1+z)^4} = 8z - 32z^2 + 88z^3 - 192z^4 + \dots,$$

whose basic implications we discussed in Basic Remark 2.11.1. This map meets the criteria in Proposition 2.9.3 for $\Sigma^1 := \{1, \infty\}$ and $\Sigma^0 := \{0, \delta, \delta/(\delta-1)\} \subset (-\infty, 1)$, and with $f(x)$ replaced by $Q(x)f(x)$ for a suitable non-zero polynomial $Q \in \mathbf{C}[x] \smallsetminus \{0\}$ such that $Q(x)f(x)$ and $Q(x)f(x/(x-1))$ are *holomorphic* (rather than merely meromorphic) under analytic continuation along the φ_*-images of all paths in $\mathbf{D} \smallsetminus \{0\}$. For the $\mathbf{Q}(x)$-linear span $\mathcal{H}$ of the five functions (2.11.10), Proposition 2.9.3 thus gives $\varphi^*\mathcal{H} \subset \mathcal{M}(\mathbf{D})$, supplying the analyticity hypotheses for Theorem 2.5.1.

By Lemma 2.11.7, the holonomy bound (2.5.4) becomes

$$5 = m \leq \frac{\log 8 + (4G/\pi)}{\log 8 - 69/50} = 4.640395\ldots,$$

a contradiction. ☐

2.11.12. *Some mixed periods.* We give an application to irrationality of the theorem we just proved.

Lemma 2.11.13. *Define*

$$\begin{aligned} H_A(x) &:= \frac{1}{\sqrt{1-4x}} \in \mathbf{Z}[\![x]\!], \\ H_B(x) &:= \frac{1}{\sqrt{1-4x}} \int_0^x \frac{1}{1-t} \frac{1}{\sqrt{1-4t}}\, dt \in \mathbf{Q}[\![x]\!], \\ H_C(x) &:= \frac{1}{\sqrt{1-4x}} \int_0^x \frac{\log(1-t)}{t\sqrt{1-4t}}\, dt \in \mathbf{Q}[\![x]\!], \\ H_D(x) &:= \frac{1}{\sqrt{1-4x}} \int_0^x \frac{\log(1-t)}{1-t} \frac{1}{\sqrt{1-4t}}\, dt \in \mathbf{Q}[\![x]\!]. \end{aligned}$$

Then $H_A(x)$, $H_B(x)$, $H_C(x)$, and $H_D(x)$ have $|x| < 1/4$ for the convergence disc of their Taylor series, and continue as holomorphic functions along all paths in $\mathbf{P}^1 \smallsetminus \{0, 1/4, 1, \infty\}$. They have the respective denominator types 1, for H_A; $[1, 2, \ldots, n]$, for H_B and $[1, 2, \ldots, n][1, 2, \ldots, n/2]$, for H_C and H_D. A simple counterclockwise loop encircling the singularity $x = 1/4$ induces the following unipotent local monodromy operator:

$$\begin{aligned} T(H_A) &= -H_A = H_A - 2H_A, \\ T(H_B) &= H_B - 2L(1, \chi_{-3})H_A, \\ T(H_C) &= H_C + \frac{\pi^2}{9} H_A, \\ T(H_D) &= H_D - 2\left(L(1, \chi_{-3}) \log 3 - L(2, \chi_{-3})\right) H_A. \end{aligned} \tag{2.11.14}$$

We also have

$$L(1, \chi_{-3}) = \frac{\pi}{3\sqrt{3}}. \tag{2.11.15}$$

Proof. All are straightforward; we indicate the computation of the $x = 1/4$ local monodromy operator T. The first equation, $T(H_A) = -H_A$, is evident as the analytic continuation must be the unique algebraic conjugate. The equation for $T(H_B)$ follows from (2.11.15) and the closed form integration evaluation (and integration by parts using $\arctan(1/\sqrt{3}) = \pi/6$)

$$H_B(x) = \frac{\pi}{3\sqrt{3}} H_A(x) - \frac{2}{\sqrt{3}} \frac{\arctan\sqrt{\frac{1-4x}{3}}}{\sqrt{1-4x}}, \tag{2.11.16}$$

where the second term is regular at $x = 1/4$. In general, just like in Basic Remark 2.10.1, for any meromorphic function $f(x)$ around $x = 1/4$ we have the meromorphy of $\frac{1}{\sqrt{1-4x}}\int_{1/4}^{x}\frac{f(t)}{\sqrt{1-4t}}\,dt$ near $x = 1/4$ (with both factors switching signs under the monodromy operator T). Consequently, the $x = 1/4$ monodromy operator T acts by

$$\begin{aligned}
T(H_A) &= -H_A = H_A - 2H_A,\\
T(H_B) &= H_B - 2H_A\int_0^{1/4}\frac{1}{1-t}\frac{1}{\sqrt{1-4t}}\,dt,\\
T(H_C) &= H_C - 2H_A\int_0^{1/4}\frac{\log(1-t)}{t\sqrt{1-4t}}\,dt,\\
T(H_D) &= H_D - 2H_A\int_0^{1/4}\frac{\log(1-t)}{1-t}\frac{1}{\sqrt{1-4t}}\,dt.
\end{aligned}$$

Fairly straightforward integrations reveal the holonomic coefficients

$$\int_0^{1/4}\frac{1}{1-t}\frac{1}{\sqrt{1-4t}}\,dt = \frac{\pi}{3\sqrt{3}} = L(1,\chi_{-3}),$$

reaffirming (2.11.16), and

$$\int_0^{1/4}\frac{\log(1-t)}{t\sqrt{1-4t}}\,dt = -\frac{\pi^2}{18}.$$

Lastly, an only slightly more involved integration — or a computing package — leads to the evaluation of

$$V(x) := \int\frac{\log(1-x)}{1-x}\frac{1}{\sqrt{1-4x}}\,dx$$

as

$$-\frac{2i}{\sqrt{3}}\left(\arctan\sqrt{\frac{1-4x}{3}}\left(\arctan\sqrt{\frac{1-4x}{3}} - i\left(\log\frac{4(1-x)}{\left(1+\sqrt{\frac{1-4x}{3}}\right)^2}\right)\right)\right)$$
$$+\frac{2i}{\sqrt{3}}\left(\mathrm{Li}_2\left(\frac{\sqrt{\frac{4x-1}{3}}-1}{\sqrt{\frac{4x-1}{3}}+1}\right)\right),$$

whereupon the familiar formulas

$$\begin{aligned}
\arctan\frac{1}{\sqrt{3}} &= \frac{\pi}{6},\\
L(1,\chi_{-3}) &= \frac{\pi}{3\sqrt{3}},\\
\mathrm{Li}_2(-1) &= -\frac{\pi^2}{12},\\
\mathrm{Li}_2\left(e^{2\pi i/3}\right) &= -\frac{\pi^2}{18} + i\frac{\sqrt{3}}{2}L(2,\chi_{-3})
\end{aligned}$$

straightforwardly evaluate the requisite holonomic coefficient

$$\int_0^{1/4}\frac{\log(1-t)}{1-t}\frac{1}{\sqrt{1-4t}}\,dt = V(1/4) - V(0) = L(1,\chi_{-3})\log 3 - L(2,\chi_{-3}).$$

The lemma follows from these integral evaluations. The special value (2.11.15), which we used in the preceding derivation, is none other than the Dirichlet class number formula for the complex quadratic field $\mathbf{Q}(\sqrt{-3})$. □

From part $(*)$ that we already proved in Theorem 2.8.4 (assuming Theorem 2.5.1, whose treatment is in § 7), we can thus readily derive a $\mathbf{Q}$-linear independence result out of the circumstance that the $x = 1/4$ local monodromy operator T simultaneously transforms H_A, H_B, H_C, and H_D by a scalar multiple of the common function H_A. The linear independence thus sifting through is for the holonomic coefficients in these monodromies:

Theorem 2.11.17. *The four periods*

$$1, \quad \frac{\pi}{\sqrt{3}}, \quad \pi^2, \quad 3L(2,\chi_{-3}) - \frac{\pi}{\sqrt{3}}\log 3$$

are $\mathbf{Q}$-linearly independent.

In particular, the Mahler measure

$$m\left(\frac{(1+x+y)^4}{3}\right) = \int_0^1 \int_0^1 \log\left|\frac{(1+e^{2\pi i s}+e^{2\pi i t})^4}{3}\right| ds\, dt \notin \mathbf{Q} \tag{2.11.18}$$

is irrational.

Proof. By Lemma 2.11.13, a $\mathbf{C}$-linear combination

$$f(x) = aH_A(x) + bH_B(x) + cH_C(x) + dH_D(x) \tag{2.11.19}$$

overconverges at the singularity $x = 1/4$ if and only if

$$a + b\frac{\pi}{3\sqrt{3}} - c\frac{\pi^2}{18} + d\left(\frac{\pi}{3\sqrt{3}}\log 3 - L(2,\chi_{-3})\right) = 0.$$

If this relation held with some nonzero integer vector $(a,b,c,d) \in \mathbf{Z}^4 \smallsetminus \{(0,0,0,0)\}$, the combination (2.11.19) would have had all the requirements of Theorem 2.8.4 with $\delta := 1/4$. Yet, clearly, $f(x)$ does not vary holonomically on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$, only on $\mathbf{P}^1 \smallsetminus \{0,\delta,1,\infty\} = \mathbf{P}^1 \smallsetminus \{0,1/4,1,\infty\}$.

The irrationality of the Mahler measure (2.11.18) follows immediately by Smyth's formula (1.1.2), which we can rewrite as

$$\begin{aligned} m\left(\frac{(1+x+y)^4}{3}\right) &= 4m(1+x+y) - \log 3 \\ &= \frac{3\sqrt{3}}{\pi}L(2,\chi_{-3}) - \log 3 = \frac{3L(2,\chi_{-2}) - \frac{\pi}{\sqrt{3}}\log 3}{\pi/\sqrt{3}}. \end{aligned}$$

This concludes the proof assuming Theorem 2.5.1, which we already proved to imply the requisite part $(*)$ of Theorem 2.8.4.

Theorem 2.5.1 will be proved in § 7, and the full Theorem 2.8.4 (which we did not need in the preceding argument) will be completed in § 6.8. □

2.12. **How we prove holonomy bounds.** We distinguish three principal steps:

(i) Setting up an auxiliary polynomials module $(Q_1,\dots,Q_m)$, by which we consider auxiliary functions such as $F := \sum_{i=1}^m Q_i f_i$ or its multivariable generalizations.

(ii) Arranging a Dirichlet box principle or Thue–Siegel lemma for the unknown coefficients of the auxiliary polynomials Q_i to have the associated function F vanish to a high order at $x = 0$.

(iii) Performing a Diophantine analysis of the lowest order coefficient of the auxiliary function F.

In especially favorable circumstances such as Hermite's approximants to the exponential function [Her1874, Her1893] or the ensuing approximants to the logarithm and binomial functions [Chu79, Chu83b], step (ii) is replaced by an explicit construction of the requisite polynomials Q_i. Some simplest examples are discussed in § 3.3. Such constructions, in the rare occasions that they are possible, usually lead to stronger quantitative results than (ii). For our intricate applications, however, as well as for the abstract theorems, some form of the Dirichlet box principle is essential.

The simplest arrangement, which already obtains *some* (rather poor) holonomy bound on the maximal number m of $\mathbf{Q}(x)$-linearly independent functions, is the following. A commonly used corollary of Siegel's lemma [BG06, Lemma 2.9.1] states that for a linear homogeneous system of M linear equations in which the coefficients are rational integers of absolute values bounded exponentially in a parameter α, while the number N of free variables is no less than twice the number of equations to be solved ($N \geq 2M$), there exist solutions whose components are rational integers, not all zero, and with absolute values bounded exponentially in $\max(\alpha, \log N)$. (Cramer's formula constructs explicitly a nonzero solution of the linear system as soon as the number of free variables strictly exceeds the number of equations; but the determinantal expression of this solution gives in general a bound which is exponential in $M\alpha$ rather than α; in our setting with $M \asymp \alpha$, this means that the Cramer solution is bounded exponentially in α^2 rather than α. For Hermite–Padé approximants to holonomic functions this is not a methodological limitation but actually the correct size in a majority of naturally occurring cases; cf. [BC97b] for a complete study of the algebraic case.)

Following Thue [Thu77, § 11], we can improve the upper bound on the solution of the linear system, from exponential in α to asymptotically subexponential in α, by using $N = (1+C)M$ free variables for a large constant C. Siegel's lemma then supplies nontrivial solutions in rational integers bounded in magnitude by $\exp\big(O(\alpha/C)\big)$; this becomes subexponential in the asymptotic where $C \to \infty$ after $\alpha \to \infty$. Hence, if we have a $\mathbf{Q}(x)$-linearly independent set $f_1, \ldots, f_m$ with denominators of the type $A^{n+1}[1, \ldots, bn]^{\sigma}$ and with m sufficiently big with regard to A, b, σ, and the smallest convergence radius of an $f_i(x)$ (this is ultimately handled in Lemma 6.2.6, in a high-dimensional setting that we will need for proving our refined bounds), Siegel's lemma guarantees the existence of a *nonzero* auxiliary function

$$F(x) := \sum_{i=1}^{m} Q_i(x) f_i(x) = \beta x^n + O(x^{n+1}) \in \mathbf{Q}[\![x]\!], \quad \beta \in \mathbf{Q}^{\times}$$

that vanishes to some high order n at $x = 0$, all the while involving integer polynomials $Q_i \in \mathbf{Z}[x]$ whose degrees and coefficients, taken on the logarithmic scale,[10] are smaller than an arbitrary desired linear rate cn in the vanishing order n.

But the meaning of "an arbitrary desired linear rate cn" is that an arbitrarily small $c > 0$ is attainable when the number m of independent functions $f_i(x)$ is supposed correspondingly large: giving a combined number of as many as $N =$

[10]This means that all these polynomials have degrees smaller than cn and rational integer coefficients with absolute values $\ll e^{cn}$.

mD undetermined coefficients for the auxiliary polynomial m-tuple $(Q_1, \ldots, Q_m) \in \mathbf{Z}[x]^{\oplus m}_{\deg < D}$. Making this quantitative will ultimately read into holonomy bounds such as (2.5.4). To explain where those derive from, observe that if the functions $f_i(x)$ are of the denominator type $[1, \ldots, b_1 n] \cdots [1, \ldots, b_r n]$, then since the auxiliary $Q_i(x)$ have integer coefficients, the lowest order coefficient $\beta \in \mathbf{Q}^\times$ is some *nonzero* rational number of this denominator, hence

$$|\beta| \geq \frac{1}{[1, \ldots, b_1 n] \cdots [1, \ldots, b_r n]}. \tag{2.12.1}$$

By the prime number theorem, this gives a Diophantine lower bound by

$$e^{-(b_1 + \ldots + b_r)n + o(n)}$$

on that leading coefficient. Now suppose we have a holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ of derivative $|\varphi'(0)| > e^{b_1 + \ldots + b_r}$ and turning all $f_i(\varphi(z)) \in \mathbf{C}[\![z]\!]$ holomorphic (*convergent*) in a neighborhood of the closed unit disc $z \in \overline{\mathbf{D}}$. Then $G(z) := z^{-n} F(\varphi(z)) = \varphi'(0)^n \beta + O(z)$ is a holomorphic function in a neighborhood of the closed unit disc, but taking an exponentially large value

$$|G(0)| = |\varphi'(0)|^n |\beta| \geq \exp\left(\left(\log|\varphi'(0)| - \sum_{h=1}^{r} b_h \right) n + o(n) \right) \tag{2.12.2}$$

at the center $z = 0$ of that disc. Yet, since by construction the degrees of the polynomials $Q_i(x)$ are smaller than cn while their coefficients are smaller than e^{cn}, we know in this construction that on the unit circle $\mathbf{T}$ the holomorphic function $G(z)$ has the pointwise upper bound

$$\sup_{\mathbf{T}} |G| \leq \exp\left(O\left(c \left(\max_{\mathbf{T}} \log|\varphi| \right) \cdot n \right) \right). \tag{2.12.3}$$

Since we can make the coefficient $c > 0$ arbitrarily small upon assuming m to be correspondingly big, but the maximum principle for holomorphic functions restrains the left-hand side of (2.12.2) to be not greater than the left-hand side of (2.12.3), our assumption of the positive rate in the lower bound (2.12.2) sets an upper limitation on the maximal number m of our $\mathbf{Q}(x)$-linearly independent functions $f_i(x)$. This dimension bound only depends on the holomorphic mapping φ and on the positive difference $\log|\varphi'(0)| - \sum_{h=1}^{r} b_h$ that occurred through (2.12.2). We call it an *arithmetic holonomy bound* due to the Diophantine way it was proved.

For simplicity of this sketch, we assumed the $f_i(\varphi(z))$ to be holomorphic rather than meromorphic functions on a neighborhood of the closed unit disc. The general meromorphic case is handled in exactly the same way just by changing the definition of the holomorphic function $G(z)$ to $G(z) := h(z) z^{-n} F(\varphi(z))$, where $h \in \mathcal{O}(\overline{\mathbf{D}})$ is a holomorphic function on a neighborhood of the closed unit disc that has $h(0) = 1$ and all $h(z) f_i(\varphi(z))$ simultaneously holomorphic on that disc.

In particular, this sketch proves André's holonomicity criterion (Corollary 2.6.1), for by the chain rule, the $\mathbf{Q}(x)$-linear span of all $f(x)$ in Corollary 2.6.1 is closed under the derivation d/dx. This is how holonomy arises out of finiteness theorems.

2.13. **Refined methods.** This subsection is a deeper and more technical introduction than the rest of § 2, and it serves as a more detailed summary of the ideas in the proofs of our holonomy bounds. It is not strictly required for the logic of these proofs. The reader might therefore opt to skip any part in the following, and refer back as needed later.

The rudimentary proof method we just described in § 2.12 is completely standard in the subject of Diophantine analysis. It is referred to as Gelfond's method in the works of Dèbes [Dèb86] and André [And89, § VIII.3], and found its first applications to arithmetic algebraization in the trailblazing work [CC85b, CC85c] of David and Gregory Chudnovsky. Our Appendix B refines these ideas through the prism of Perelli and Zannier's work [PZ84] to re-derive the bound (2.2.3) in our context, including the $e \rightsquigarrow 2$ coefficient reduction by a single-variable analysis. As mentioned in § 2.3, for our applications to irrationality, we have two alternative lines of holonomy bounds: one via high-dimensional techniques (Theorem 6.0.2, which implies (2.5.5)), and the other via the single variable slopes method (Theorem 7.0.1, which implies Theorem 2.5.1, and its strengthening Theorem 7.1.6). We now discuss what can be improved in the preceding scheme to obtain these two lines of refined results separately. We begin with the ideas of the proof of Theorem 6.0.2, based on Diophantine approximation in several variables. The basic idea can be summarized by saying that our multivariable evaluation module will lose none of the simplicity of the essentially one-dimensional features similar to § B, yet it also has all the added flexibility of the Law of Large Numbers inherent in any Diophantine approximation scheme with $d \to \infty$ variables.

2.13.1. *The possible vanishing orders.* We can formulate step (i) of the preceding scheme differently. We do this just as easily in a multivariable framework with $\mathbf{x} := (x_1, \ldots, x_d)$, which as we will see is ultimately advantageous for the proofs upon working with the d-th Cartesian power of the single-variable evaluation module. Given

- a $\mathbf{Q}(\mathbf{x})$-linearly independent set $\{f_{\mathbf{i}}(\mathbf{x})\}_{\mathbf{i}\in I}$ of $\mathbf{Q}[\![\mathbf{x}]\!]$ formal power series, to be indexed by a finite set I which for our purposes will be taken a subset $I \subset \{1, \ldots, m\}^d$,
- and a bounded Lebesgue-measurable subset $\Omega \subset [0,\infty)^d$,

we can express the preceding argument by introducing a parameter D and taking $(Q_1, \ldots, Q_m)$, or $(Q_{\mathbf{i}})_{\mathbf{i}\in I}$ in this generality, to range from the *auxiliary polynomials module*

$$E^I_{D,\Omega} := \mathrm{Span}_{\mathbf{Z}} \left\{ \mathbf{x}^{\mathbf{k}} \,:\, \mathbf{k} \in (D\cdot\Omega) \cap \mathbf{Z}^d \right\}^{\oplus I},$$

a free $\mathbf{Z}$-module of rank $R^I_{D,\Omega} = (1+o(1))(\#I)\,\mathrm{vol}(\Omega)D^d$. For the original case of $d = 1$, $I = \{1, \ldots, m\}$, and $\Omega = [0, 1)$, we simplify the notation to $E_D = \mathbf{Z}[x]^{\oplus m}_{<D}$, of rank $R^{\{1,\ldots,m\}}_{D,[0,1)} = mD$.

The $\mathbf{Q}(\mathbf{x})$-linear independence condition on the $f_{\mathbf{i}}(\mathbf{x})$ means exactly that, for all D and Ω, the *evaluation homomorphism*

$$\psi_D : E^I_{D,\Omega} \hookrightarrow \mathbf{Q}[\![\mathbf{x}]\!], \qquad (Q_{\mathbf{i}})_{\mathbf{i}\in I} \mapsto \sum_{\mathbf{i}\in I} Q_{\mathbf{i}} f_{\mathbf{i}} \in \mathbf{Q}[\![\mathbf{x}]\!],$$

is injective. (We drop Ω and I from the notation of ψ_D, as they will ultimately be considered fixed throughout the procedure, whereas D will be the first asymptotic parameter to be let $\to \infty$.)

In the outline § 2.12, we considered some power series $F(x) = \beta x^n + O(x^{n+1})$ from the range of this evaluation map (for $d = 1$) that vanished at $x = 0$ to the exact order n. But the possible leading order exponents $\mathbf{n} \in \mathbf{N}^d$ in any $F = \sum_{\mathbf{i}} Q_{\mathbf{i}} f_{\mathbf{i}} \in E^I_{D,\Omega}$ take up exactly $R^I_{D,\Omega} = \dim_{\mathbf{Q}}(E^I_{D,\Omega} \otimes \mathbf{Q})$ possibilities that depend only on the evaluation module (E_D, ψ_D), and not on the specific element $(Q_{\mathbf{i}}) \in E^I_{D,\Omega}$. These *(vanishing) filtration jumps*[11] form a size-$R^I_{D,\Omega}$ subset of $\mathbf{N}^d$, which we formally define in § 3.1.3.

For our final results in this paper, we ultimately only consider single variable ODEs. The high-dimensional modules $E^I_{D,\Omega}$ arise from involving the d-fold Cartesian power

$$E^{\{1,\dots,m\}^d}_{D,[0,1)^d} = E_D \times \cdots \times E_D, \quad f_{\mathbf{i}}(\mathbf{x}) := \prod_{s=1}^d f_{i_s}(x_s), \quad \text{of rank } (mD)^d, \tag{2.13.2}$$

of the univariate module $E_D = E^{\{1,\dots,m\}}_{D,[0,1)}$ generated by the functions $f_1, \dots, f_m \in \mathbf{Q}[\![x]\!]$ of Theorem 2.5.1, and their suitable submodules — this is the idea of measure concentration in the $d \to \infty$ limit — given by restriction to statistically preponderant subsets $\Omega \subset [0,1)^d$ and $I \subset \{1, \dots, m\}^d$.

A basic idea for our new developments here over the results in [CDT21, § 2] is a simple lemma (Corollary 3.1.11) about the commutation in the formations of Cartesian products of evaluation modules and the sets of vanishing filtration jumps. Concretely, the $(mD)^d$ filtration jumps of the evaluation module (2.13.2) are at a Cartesian power set of the form S^d for some $S \subset \mathbf{N}$ with $\#S = mD$.

2.13.3. *Methods from differential algebra and functional bad approximability.* In our proofs of the general holonomy bounds, we use a more precise information on the vanishing filtration jumps. This takes on the role of the "zero estimates" in the traditional transcendence theory proofs. In our context, the latter can be seen as functional analogs of the Schmidt Subspace theorem on bad approximability. (See, indeed, [Wan04] for the case of algebraic functions.) Easier but cruder versions — analogous rather to Liouville's Diophantine inequality for differential algebra — include the prototypical Shidlovsky lemma [Shi89, § 3.5, Lemma 8] from the historical proof [Shi59] of the Siegel–Shidlovsky theorem on special values of E-functions, with its multitude of effectively computable variations [Chu80, § 11], [BB85, BCY04], [Ber12, § 2] available in the literature. The general bad approximability theorem was known as Kolchin's problem ([Kol59], see Problem 3.2.7), before it was proven independently by David and Gregory Chudnovsky [CC83] and Osgood [Osg85], for the essential case of holonomic $f_1, \dots, f_m$. Its statement amounts to saying that the vanishing filtration jumps set S is close to the generic jumps $\{0, 1, \dots, mD-1\}$, in the sense that

$$S \subset \big\{0, 1, \dots, mD + o_{f_1,\dots,f_m}(D)\big\}, \qquad \#S = mD.$$

In an asymptotic sense, this almost determines the vanishing filtration jumps for all the holonomic evaluation modules of relevance to our paper: those being the modules $E^I_{D,\Omega}$ with $\Omega \subset [0,1)^d$ of $\mathrm{vol}(\Omega) = 1 - o_{d\to\infty}(1)$; $I \subset \{1, \dots, m\}^d$ with $\#I = m^d - o_{d\to\infty}(m^d)$, and holonomic $f_1, \dots, f_m$. These improvements are discussed in § 3.2.

Technically, for the qualitative linear independence proofs of Theorem A and Theorem C (up to replacing the numerical threshold 10^{-6} by a smaller absolute constant), it is actually possible to avoid all recourse to this differential algebra material § 3.2. It is however an unnecessarily convoluted route to insist on; moreover, some version of the theorems collected in § 3.2 is indispensable in pursuing

[11] They may as well be termed the *successive minima* of the evaluation module, as in [Ber99] taking an inspiration from the Weierstrass gaps on algebraic curves.

any quantitative refinements to Diophantine measures of linear independence. We choose to use functional bad approximability in our main proofs as well, for the holonomy bounds in §§ 6–8, as that allows for cleaner arguments, and is actually (as far as we are aware) necessary for most of the general — qualitative! — holonomy bounds in the clean structural form in which we have stated them. We do observe, however, that such structural necessities do not concern the main $|\varphi'(0)| > e^{\sigma_m}$ case of Theorem 2.5.1 itself, which does admit clean proofs not relying on any functional bad approximability theorems. (Remark 6.0.16 shows that the $|\varphi'(0)| > e^{\tau(\mathbf{b})}$ case *must* make some special use of the ODE.) All this is discussed in § 7.7. The reader may compare the situation with the simpler § B, where no special information on the vanishing filtration jumps is relevant to the proof of the qualitative holonomy bound (B.0.1).

2.13.4. *Multiple variables unlock the Law of Large Numbers.* We next discuss how, in the $d \to \infty$ asymptotic modeled by independent and identically distributed random variables, we can exploit the full-measure subsets $\Omega \subset [0,1)^d$ and $I \subset \{1,\ldots,m\}^d$. Historically, Diophantine approximation by multiple variables was the key to refining Liouville's bad approximability theorem $|\alpha - p/q| \gg q^{-[\mathbf{Q}(\alpha):\mathbf{Q}]}$ to Roth's "best-possible" bad approximability measure $|\alpha - p/q| \gg_\varepsilon q^{-2-\varepsilon}$ (when the target α is algebraic and irrational). The purpose of the scheme[12] is to make the maximum use of the free parameters count in the application of Siegel's lemma. Having a multivariable auxiliary function $F(x_1,\ldots,x_d)$ vanish to a high $(D_1,\ldots,D_d)$-weighted order $\geq \xi d$ at a point $(0,\ldots,0)$ means to vanish all monomials $\mathbf{x}^{\mathbf{n}} := x_1^{n_1}\cdots x_d^{n_d}$ with $n_1/D_1 + \ldots + n_d/D_d \geq \xi d$. But as $d \to \infty$ and $D_i \to \infty$ with $t_i := n_i/D_i \in [0,1]$, the Law of Large Numbers (in Chernoff's form) for the sum $\sum_{i=1}^d t_i \approx d/2$ of $d \to \infty$ uniform and identically distributed random variables $t_i \in [0,1]$ shows that, with an $1 - \exp(O(-d\varepsilon^2))$ probability, $\xi = \frac{1}{2} - \varepsilon$ is the correct reasonable weighted vanishing order to attain by the parameter count in the Thue–Siegel lemma. (To contrast, the single variable construction only reaches the Liouville-strength vanishing order coefficient $\xi = 1/[\mathbf{Q}(\alpha):\mathbf{Q}]$, and the two-variables construction only reaches a vanishing order coefficient of about $\xi = 1/(2\sqrt{[\mathbf{Q}(\alpha):\mathbf{Q}]})$, giving the exponent in Siegel's sub-Liouville theorem [Sie1921].)

Further work of Wirsing [Wir71, see § 4.2], aimed at correcting Roth's Corrigendum in [Rot55] regarding approximation of an algebraic number target by algebraic number approximants of a fixed degree over $\mathbf{Q}$, pivoted around a refinement of the above Law of Large numbers, the *measure concentration property of the high-dimensional hypercube* $[0,1]^d$, which states that not only $\frac{1}{d}\sum_{i=1}^d t_i$ converges in probability to the expectation $\mathbf{E}[t] = \int_0^1 t\,dt = \frac{1}{2}$ as $d \to \infty$, but further and more precisely, that with high asymptotic probability as $d \to \infty$, the random vector $(t_1,\ldots,t_d) \in [0,1]^d$ has uniformly distributed components. This has a precise meaning in our Theorem 4.2.1 below refining [Wir71, Lemma 13]: the ε-high discrepancy set (see Definition 4.1.1)

$$B_\varepsilon^d := \left\{\mathbf{t} \in [0,1)^d \,:\, \exists [a,b) \subset [0,1),\ \left|(b-a) - \frac{1}{d}\#\{i \,:\, t_i \in I\}\right| \geq \varepsilon\right\}$$

[12]Found by Siegel, and attempted with partial success by Schneider [Sch36] prior to Roth's work [Rot55].

has d-dimensional Lebesgue measure $\mathrm{vol}(B_\varepsilon^d) \leq 100 \exp\left(-\varepsilon^4 d/300\right)$. This is ultimately the statistical property behind the rearrangement integral in our bound

$$m \leq \frac{\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t})|)^* \, dt}{\log|\varphi'(0)| - \tau(\mathbf{b})} \tag{2.13.5}$$

discussed in § 2.4; this bound is a special case of Theorem 6.0.2 with $\mathbf{e} = \mathbf{0}, l = 0, \varphi_0 = \varphi$.

We review in § 4 the topic of measure concentration and large statistical deviations. These ideas are used not only to sift through the subsets $\Omega \subset [0,1)^d$ and $I \subset \{1, \ldots, m\}^d$ in the make-up of the evaluation module $E^I_{D,\Omega}$, but also to usefully limit the shape of the monomials $\mathbf{x}^{\mathbf{n}}$ from the leading order jet of the d-variate auxiliary function $F \in \sum_{\mathbf{i}} Q_{\mathbf{i}} f_{\mathbf{i}} \in \mathbf{Q}[\![\mathbf{x}]\!] \setminus \{0\}$. The former leads to the rearrangement integral; the latter two lead, in particular, to the refined denominators rate $\tau(\mathbf{b})$. We discuss in § 2.13.6 the mechanism for both these improvements. As we have mentioned in § 2.3, for the case of basic denominator types as in Theorem 2.5.1 (as well as in all the other holonomy bounds in §§ 6–7), the exact same denominator saving comes through also by the single variable method (hence no measure concentration) of § 7; while the best general denominator term comes through in Theorem 8.0.1, again by measure concentration. Although the proofs of Theorems 6.0.2 and 8.0.1 are described in different languages (one via the Thue–Siegel Lemma, the other via Bost's slopes method), the ideas on treating denominators behind both proofs are the same, as is the scope for the further improvements in the denominators aspect.

In all three proofs of our holonomy bounds in [CDT21, § 2], we used $d \to \infty$ for its automatic improvement of the Dirichlet exponent — namely, if M is the number of equations and N is the number of parameters, then $M/N = o_{d\to\infty}(1)$, and hence the Dirichlet exponent $M/(N-M)$ is also $o_{d\to\infty}(1)$. In this paper, this aspect is shown again in (6.3.4), but for this particular point, $d \to \infty$ is used only as a methodological feature of working with the most traditional form of the Thue–Siegel lemma. (Appendix B explains how we could bypass the auxiliary coefficients size while sticking to the single variable module E_D, in a form similar to the slopes method treatment in § 7.) The input from measure concentration is by far the more essential use of the high dimensions.

The fine improvements in the numerator and denominator of the fraction (2.2.3) are however only relevant insofar that they also come with an $e \rightsquigarrow 2$ overall coefficient reduction. We discuss next how this is achieved by exploiting, in the Dirichlet box principle, Lemma 3.1.11 on the Cartesian power structure of the vanishing filtration jumps, in the sense described in § 2.13.1. This point will also clarify the employment of the functional bad approximability results that we mentioned in § 2.13.3.

2.13.6. *The high-dimensional parameter count.* At the outset, to have an auxiliary function $F := \sum Q_{\mathbf{i}} f_{\mathbf{i}}$ in the range of the general evaluation module $(E^I_{\Omega,D}, \psi_D)$ to vanish to an order at least α at $\mathbf{x} = \mathbf{0}$, involves solving $\binom{\alpha+d}{d} \sim \alpha^d/d!$ linear equations in the $\sim \mathrm{vol}(\Omega)(mD)^d$ unknown coefficients of the polynomials $Q_{\mathbf{i}}$. By Stirling's asymptotic $d! = d^d/e^{d-o(d)}$, the maximal attainable vanishing order in the high-dimensional asymptotic $d \to \infty$ appears to be $\alpha \sim mdD/e$. This was why in [CDT21, § 2] we have the coefficient e in the holonomy bound that we established there with the hypercube choice $\Omega := [0,1)^d$. Had we used instead the simplex choice

$$\Omega := \{(t_1,\dots,t_d) \in [0,1] \,:\, t_1+\ldots+t_d < 1\}\,, \qquad \mathrm{vol}(\Omega) = 1/d!,$$

we would have entertained an asymptotic vanishing order as high as $\alpha \sim mD$ (without the number e entering in as a coefficient); but in this case the functions $Q_{\mathbf{i}}(\varphi(z_1),\dots,\varphi(z_d))$ would be far too big on the unit polycircle $\mathbf{z} = (z_1,\dots,z_d) \in \mathbf{T}^d$. In such an approach, we would have only obtained an inadequately big holonomy bound with (in the context) an exponentially larger numerator such as $\sup_{\mathbf{T}} \log|\varphi|$, instead of the Nevanlinna growth characteristic $T(\varphi) = \int_{\mathbf{T}} \log^+|\varphi|\,\mu_{\mathrm{Haar}}$. In the present paper, for a similar reason, we still use the hypercube shape $\Omega = [0,1)^d$, or more precisely, its measure-concentrated subsets $\Omega = P_\epsilon^d := [0,1)^d \setminus B_\epsilon^d$ for the auxiliary monomials exponents range. As we discussed above, we do rely on these statistically preponderant parts of the high-dimensional hypercube in order to get the refined growth integral (2.4.2); moreover, we will explain how we use these statistics to control the shape of lowest order terms in Siegel Lemma construction in order to obtain the denominators counterpart (2.5.6) of the refined growth integral.

The Cartesian power situation is special for enforcing, as discussed in § 2.13.1, a Cartesian power structure (Corollary 3.1.11) on the vanishing filtration jumps vectors $\subset \mathbf{N}^d$ of $E_{[0,1)^d,D}$. These are in turn brought to exploit a certain automatic vanishing of many of the coefficients of the sought-for auxiliary function F. The simplest instance of this automatic vanishing is showcased in § B.2. Instead of directly solving for the vanishing of all the low-degree monomials of F (which we definitely need for the maximum modulus principle step when we carry out the higher-dimensional extension of step (iii) of § 2.12), we set up the Thue–Siegel lemma differently by focusing on the mD filtration jumps

$$0 \leq u(1) < u(2) < \cdots < u(mD)$$

of the single-variable evaluation module E_D. In the single-variable situation the procedure simply reduces to setting to zero the $x^{u(p)}$ coefficient $\beta_{u(p)} = 0$ of $F(x)$ for $p = 1,\dots,mD$. In general, for any subset $T \subset [0,mD]^d$, we write

$$u(T) := \big\{(u(s_1),\dots,u(s_d)) \,:\, (s_1,\dots,s_d) \in T \cap \mathbf{N}_{>0}^d\big\}.$$

Then, in the measure-concentrated submodule $E_{D,\Omega}^I \subset E_{D,[0,1)^d}^{\{1,\dots,mD\}^d}$ with $\mathrm{vol}(\Omega) > 1 - 100\exp\left(-\epsilon^4 d/300\right)$, we use our $(mD)^{d-o(d)}$ degrees of freedom in the auxiliary polynomials coefficients to construct a nonzero $F(\mathbf{x}) = \sum_{\mathbf{i}} Q_{\mathbf{i}}(\mathbf{x}) f_{\mathbf{i}}(\mathbf{x}) = \sum \beta_{\mathbf{n}} \mathbf{x}^{\mathbf{n}}$ with all auxiliary polynomials $Q_{\mathbf{i}}$ having integer coefficients bounded by $e^{\epsilon d D}$ in absolute value, and in which (with a sufficiently small $\delta \in (0,\epsilon)$)

$$\beta_{\mathbf{n}} = 0 \quad \text{for all} \quad \mathbf{n} \in u\big([0,(m-\delta)D]^d\big) \cup u\big((m+\delta)D \cdot B_\epsilon^d\big), \tag{2.13.7}$$

provided $\delta \in (0,\epsilon)$ is small enough to have $(m-\delta)/(m+\delta) > \exp\left(-\epsilon^4/400\right)$. For Theorem 2.5.1 when $|\varphi'(0)| > e^{\sigma_m}$, and for some further forms of our bounds that are discussed in § 7.7 (which do cover, in particular, the ultimate application to Theorem A), it is technically possible to devise a proof directly out of this construction, and without appeal to the ideas of § 2.13.3.

In any event, for our practical purposes in this paper, if the reader would like to further simplify the essential mental picture, it would be very reasonable to imagine at this point that the filtration jumps are as simple as possible, namely given by $u(i) = i-1$. Such is for example the case with the classical Hermite–Padé systems that we discuss in § 3.3. The tenor of the functional bad approximability theorems of § 3.2 is that, for the purposes of many applications including ours, such an assumption is not far from being satisfied: the Chudnovsky–Osgood theorem, as we formulated in § 2.13.3, can be stated as the upper bound $u(mD) \le (m+\varepsilon)D + C(\varepsilon)$, which is at most $(m+\epsilon)D$ for $D \gg 1$ if we assume $\varepsilon < \delta < \epsilon$. We observe as a statistical effect that the difference between $u(i)$ and its lower bound $i-1$ becomes negligible in the asymptotic analysis of $D \to \infty$ followed by $\epsilon \to 0$. Our gateway to the functional bad approximability theorems is through André's holonomicity criterion (Corollary 2.6.1; unless, as in the applications, the f_i are *a priori* given holonomic), whose proof was outlined in § 2.12 and laid out in full in § B. With the Chudnovsky–Osgood theorem, the previous construction reduces simply to attaining

$$\beta_{\mathbf{n}} = 0 \quad \text{for all} \quad \mathbf{n} \in [0, (m-\delta)D]^d \cup (m+\delta)D \cdot B_\epsilon^d. \tag{2.13.8}$$

Whichever the approach, the routine for step (iii) of 2.12 is to examine the possibilities for a lowest-order nonzero coefficient $\beta := \beta_{\mathbf{n}} \neq 0$. On the one hand, as we discussed above, the Cartesian structure restrains $\mathbf{n}$ to be of the form

$$\mathbf{n} = (u(p_1), u(p_2), \ldots, u(p_d)), \qquad \text{for some } p_1, \ldots, p_d \in \{1, \ldots, mD\}.$$

Since our Thue–Siegel lemma construction disposed of all the multi-indices $(p_1, \ldots, p_d)$ in $(mD+\delta) \cdot B_\epsilon^d$ lying in the ϵ-high discrepancy part of the hypercube, the above tuple $(p_1, \ldots, p_d)$ must belong to the complementary part $(mD+\delta) \cdot P_\epsilon^d$ of the statistically typical points. Heuristically speaking, the components $n_j = u(p_j)$ of each lowest-order exponent vector $\mathbf{n}$ in the Taylor series of $F \in \mathbf{Q}[\![\mathbf{x}]\!] \smallsetminus \{0\}$ are close — as $d \to \infty$ followed by $\epsilon \to 0$ — to some ordering of the set

$$\{u(\lfloor mD/d \rfloor), u(\lfloor 2mD/d \rfloor), \ldots, u(\lfloor dmD/d \rfloor)\}.$$

In particular, the vanishing order in this auxiliary construction satisfies

$$\begin{aligned} \mathrm{ord}_{\mathbf{x}=\mathbf{0}} F = |\mathbf{n}| &= (1+o(1)) \sum_{j=1}^{d} u(\lfloor jmD/d \rfloor) \\ &\ge (1+o(1)) \sum_{j=1}^{d} jmD/d = (1+o(1)) mdD/2, \end{aligned}$$

a notable improvement of the asymptotic vanishing order parameter $\alpha \sim mdD/e$ in [CDT21, § 2].

This heuristic lower estimate does indeed match the accurate asymptotic formula from using the Chudnovsky–Osgood theorem and (2.13.8). The one (fundamentally minor) technical point in arguing directly from (2.13.7), for the reader who may desire additionally here to forsake the theorems in § 3.2, is that — for the discrepancy theory purposes of our proofs — the uniform distribution $\{p_1, \ldots, p_d\} \approx \{\lfloor mD/d \rfloor, \lfloor 2mD/d \rfloor, \ldots, \lfloor dmD/d \rfloor\}$ does not preserve the $\approx$ relation upon applying u to both sides. This is however irrelevant to the above outline; all that matters is that the facts that $u(i) \ge i-1$ and that $(p_1, \ldots, p_d)$ has asymptotically uniformly distributed components by themselves imply $|\mathbf{n}| \ge (1+o(1))mdD/2$.

2.13.9. *Effects on denominators.* We now discuss how to obtain the refined denominators saving in Theorem 6.0.2, and by extension, in Theorem 8.0.1. To illustrate the idea, we use the construction (2.13.8) contingent upon the Chudnovsky–Osgood theorem, and we consider the lexicographically minimal term $\beta\, \mathbf{x}^{\mathbf{n}}$ in $F(\mathbf{x})$ among all the terms of the minimal total degree $n = |\mathbf{n}|$. Thus $\mathbf{n} \in (m+\delta)D \cdot P_\epsilon^d$, and F is a $\mathbf{Q}[\mathbf{x}]$-linear combination of $f_{\mathbf{i}}$, where every $\mathbf{i} \in I$ is *balanced*, namely each $i_0 \in \{1,\dots,m\}$ occurs about d/m times among all $i_j, 1 \leq j \leq d$. The denominator of the nonzero rational number $\beta \in \mathbf{Q}^\times$ divides the lowest common multiple of the denominators of the $\mathbf{x}^{\mathbf{n}}$ terms in all formal functions from the modules $\mathbf{Q}[\mathbf{x}]f_{\mathbf{i}}$, as $\mathbf{i} \in I$ ranges over the balanced multi-index sets. The particular form of the denominators assumed in Theorem 2.5.1 — with the types of $f_1,\dots,f_m$ being "from best to worst" in this order — implies that said lowest common multiple formally agrees asymptotically with the $\mathbf{x}^{\mathbf{n}}$ coefficient of $f_{\mathbf{i}_0}$, where $\mathbf{i}_0$ is a balanced multi-index arranged in nondecreasing order. This observation yields our denominator saving term $\tau(\mathbf{b})$ as a "finite rearrangement integral" (2.5.6). In general, there is not a single particular $\mathbf{i}$ to make the asymptotic denominator of β; this "collective $\mathbf{i}_0$" is rather the formal effect of working only with the balanced $\mathbf{i}$, which — as another effect of the measure concentration[13] advantage of $d \to \infty$ — are statistically preponderant in $\{1,\dots,m\}^d$.

2.13.10. *Complex-analytic tools.* The maximum principle can be replaced by the Poisson–Jensen formula (§ 8.2.11 or [CDT21, § 2.4]) or enhanced by seeking the optimal quotient representation $\varphi = v/u$ by holomorphic functions $v, u \in \mathcal{O}(\overline{\mathbf{D}})$ with $u(0) = 1$ (§ B.3 and [CDT21, § 2.3]). But in the $d \to \infty$ asymptotic we discussed in § 2.13.6, a better holomorphic dampener than $u(z_1)^D \cdots u(z_d)^D$ to use in the multivariable analytic function $F(\varphi(z_1),\dots,\varphi(z_d))$ would be to take a suitable power of the discriminant polynomial $\prod_{1\leq i<j\leq d}(z_i - z_j)$, which is very small on the part of the torus $\mathbf{z} = (z_1,\dots,z_d) \in \mathbf{T}^d$ where the set $\{z_1,\dots,z_d\}$ has a non-small discrepancy from the uniform measure μ_{Haar} of the circle $\mathbf{T}$. This was the *ad hoc* method in [CDT21, § 2.5], which fits here the most naturally into the cross-variables integration scheme stemming from § 2.13.6, ultimately leading into the bound (2.13.5). This is our treatment in § 6.5. In § 8.2.11, we give a second treatment based on the Poisson–Jensen formula.

The further refinements that we mentioned in § 2.3 are based on the following idea. If we consider another holomorphic mapping $\psi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ also having all $\psi^* f_i \in \mathcal{M}(\mathbf{D})$, we may replace a subset of the $\varphi(z_j)$ in $F(\varphi(z_1),\dots,\varphi(z_d))$ by $\psi(z_j)$, and carry out a similar analysis thus using the combined analytic maps φ and ψ. To use φ for $j \in S_1$ and ψ for $j \in S_2$, for some partitioning $\{1,\dots,d\} = S_1 \sqcup S_2$ of the indexing set into proportionally large subsets S_1 and S_2, observe that upon taking our holomorphic dampener to be a suitable power of

$$\prod_{1\leq i<j\leq d, i,j\in S_1} (z_i - z_j) \prod_{1\leq i<j\leq d, i,j\in S_2} (z_i - z_j),$$

the main contribution to the growth of the auxiliary function pullback on $\mathbf{T}^d$ comes from the part of the torus $\mathbf{z} \in \mathbf{T}^d$ where *both* sequences $(z_j)_{j\in S_1}$ and $(z_j)_{j\in S_2}$ have a small discrepancy from the uniform distribution on the circle $\mathbf{T}$. The point is that, when we estimate the leading-order $\mathbf{x}^{\mathbf{n}}$ coefficient β by the analytic method, only the variables indexed by $j \in S_1$ use φ while the variables indexed by $j \in S_2$ use ψ. We select the partition so as to minimize the upper bound from maximum principle over $\mathbf{z} \in \mathbf{T}^d$. For a given φ, we may certainly take our second (or, repeating the procedure, our "next") map to be $\psi(z) := \varphi(rz)$ for an arbitrary $0 < r < 1$. As far

[13]Here basically amounting to the maximality of the central multinomial coefficients, cf. Lemma 6.2.4.

as φ is not univalent, we prove that depending on the size of n_j, one may choose for the variable z_j a certain optimal radius $r = r(n_j) \in (0,1)$ to obtain a strictly better estimate. This is the idea behind the improvement from the bound (2.13.5) to the full Theorem 6.0.2.

2.13.11. *A dynamic box principle or a finer Geometry of Numbers.* The basic sketch given in § 2.12 was grounded in a "static" Thue–Siegel lemma construction: finding a nonzero auxiliary function $F \in E_D$, then arguing "by extrapolation" from putting together the arithmetic and the analytic properties of the lowest-order nonzero coefficient $\beta \in \mathbf{Q}^\times$. This simple-minded procedure is insufficient for obtaining the original holonomy bound (2.1.1) by a single-variable analysis, because in the Thue–Siegel lemma of § 2.12 it is impossible to attain a small Dirichlet exponent $M/(N-M) < c$ all the while having a near-maximal vanishing order $M \approx N$. In [CDT21, § 2] we exploited the decaying Dirichlet exponent under $d \to \infty$ (in the present paper, this is the step (6.3.4)), making the issue go away in the high-dimensional analysis. As the Bost–Charles work [BC22] made it abundantly clear, it is possible to prove (2.2.3), even with the coefficient reduction $e \rightsquigarrow 2$, by one-dimensional methods once the rudimentary Thue–Siegel lemma is replaced by a sufficiently precise arrangement of the pigeonhole or Minkowski arguments. Our Appendix § B gives an essentially elementary such proof based on the dynamic box principle technique of Perelli and Zannier [PZ84, Lemma 1]. This may be also read as an introduction to Bost's slopes method framework, whose idea is very similar but cast into the language of Hermitian vector bundles over Spec $\mathbf{Z}$, and which is the content of § 7. △

We now discuss the more specific ideas for the proofs of Theorems 7.0.1, 7.1.6, and 7.1.13 via Bost's method of slopes. Common ingredients (with the simplification applied to a single variable situation) are §§ 2.13.3 and 2.13.10.

2.13.12. *Bost's slopes method with ingredients from Bost–Charles* [BC22]. We adapt the notation from § 2.13.1 to consider a filtered $\mathbf{Z}$-module E_D and an evaluation homomorphism ψ_D. (In the bulk of § 7, we opt to rather work with $x^{1-D}E_D$ as that allows for a more natural identification with the global sections of a certain ample line bundle; for simplicity here, we stick to the positive degree monomials, like we do in one of our more elementary slopes method variations in § 7.5.) We let $E_D^{(n)} \subset E_D$ to denote the n^{th} vanishing order filtration, namely the submodule consisting of those elements whose image under ψ_D vanishes to order at least n at $x = 0$. The evaluation homomorphism $\psi_D : E_D \hookrightarrow \mathbf{Q}[\![x]\!]$ then induces a set of *monomorphisms* $\psi_D^{(n)}$ on the graded quotients. Once one endows E_D with a Euclidean lattice structure, one can define an *arithmetic degree* of the underlying Hermitian vector bundle $\overline{E}_D$, and the *heights* of the evaluation maps $\psi_D^{(n)}$. (Doing this involves fixing a lattice of $\mathbf{Q}[\![x]\!]$ and endowing it with a pro-Euclidean structure. This then defines the local and global heights of $\psi_D^{(n)}$ following [Bos20, § 1.4.3]. We stick to the natural lattice choice, namely $\mathbf{Z}[\![x]\!]$ with pro-Euclidean structure induced from using $\{x^n\}_{n=0}^{N-1}$ for an orthonormal basis of each finite-dimensional quotient $\mathbf{R}[\![x]\!]/x^N\mathbf{R}[\![x]\!]$ of $\mathbf{R}[\![x]\!]$.)

Bost's slopes inequality (7.2.14) provides an upper bound on the arithmetic degree of $\overline{E}_D$ in terms of the heights of the evaluation maps $\psi_D^{(n)}$. In [Bos01, Bos04], Bost proved various algebraicity criteria in arithmetic-geometric settings similar to § 2.1. His methods combined a crude version of the global arithmetic Hilbert–Samuel formula, used as a lower bound on the arithmetic degree of $\overline{E}_D$, with local complex and p-adic analysis tools, employed to devise place-by-place upper bounds on all the local heights of all the evaluation maps $\psi_D^{(n)}$. The algebraization results then sift out from the slopes inequality under the $D \to \infty$ asymptotic. The recent work of Bost and Charles [BC22] is written (in part) under the framework of Bost's theory [Bos20] of theta invariants of infinite-dimensional Hermitian vector bundles over arithmetic curves, but one can certainly interpret the arguments in the language of the more rudimentary slopes method. We stick to the latter choice because the convexity enhancements in § 7.1 seem to be more of an analytic than a geometric nature, and we do not attempt here to include these into the theory of the theta invariant.

The main ingredients of the proofs of the bounds (2.2.4) and (2.2.5) are the arithmetic Hilbert–Samuel formula for the exact asymptotic arithmetic degree of $\overline{E}_D$, and a choice of the Euclidean structure giving rise to $\overline{E}_D$ based upon optimizing the complex analysis of the archimedean local heights of $\psi_D^{(n)}$. The latter relies on the standard tools of the subject: the Poincaré–Lelong and Poisson–Jensen formulas. One technical point in Bost and Charles's theory [BC22, § 4], needed for carrying out the arithmetic intersection number computations, is to extend the scope of the classical Arakelov theory to allow for Green functions and Hermitian metrics that are not necessarily smooth but have, in Bost and Charles's terminology, a $\mathcal{C}^{\mathrm{b}\Delta}$ regularity: a condition [BC22, Def. 4.1.1] related to using continuous Green functions locally of bounded variation. We use this framework in §§ 7, 8.

2.13.13. *Varia.* To prove Theorem 7.0.1, we use the same Euclidean norm on $E_D \otimes_{\mathbf{Z}} \mathbf{R}$ as alluded to in the final paragraph of § 2.13.12; and we adapt the same complex analytic estimates on the archimedean local heights of $\psi_D^{(n)}$. On the other hand, based on the denominator type (7.0.1), we choose a new $\mathbf{Z}$-sublattice of $E_D \otimes_{\mathbf{Z}} \mathbf{Q}$ to optimize the comparison between the arithmetic degree of $\overline{E}_D$ and the combined non-archimedean heights of all the evaluation maps $\psi_D^{(n)}$ involved in the slopes inequality (7.2.14). The latter brings out the vanishing filtration jumps sets § 2.13.1 of the evaluation module, and this is where the theorems from § 3.2 (as summarized by § 2.13.3) are used in this method also.

The extra input for Theorem 7.1.6 rests on the idea of § 2.13.10 with the multiple holomorphic maps $\varphi, \psi, \dots$ for devising sharper estimates on the various archimedean local evaluation heights $h_\infty(\psi_D^{(n)})$ in accordance with the range of n/D. We stick to the choice $\psi(z) = \varphi(rz)$ derived from a single holomorphic ambience $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$, where the convexity property in $\log r$ for various Nevanlinna-style growth characteristic functions implies that, in the multivalent case, there is always some improvement from every single intermediate radius $r \in (0,1]$. Ultimately this leads to the limiting form in Theorem 7.1.10, where the total convexity saving is presented as a dr/r integral over $r \in [0,1]$ of the square of an analog of the Ahlfors–Shimizu covering area function. Such a principle applies to a number of variations of the Nevanlinna characteristic of a meromorphic map that could be used for the principal term of the holonomy bounds, including the traditional Ahlfors–Shimizu characteristic figuring in [BC22, Prop. 5.4.5]; more significantly for us (see Ex. 8.1.16 for an illustrative comparison), it holds for the Bost–Charles characteristic that we define in § 7.1.2 (sticking for simplicity to the most basic case that we use of a holomorphic mapping $\overline{\mathbf{D}} \to \mathbf{C}$).

We remark on the other hand that once we involve this new improvement on the archimedean local evaluation height bounds from using a finite set of intermediate radii r, Bost and Charles's choice of Euclidean norm on $E_D \otimes_{\mathbf{Z}} \mathbf{R}$ is no longer the optimal in general (unless φ is univalent). We propose in Theorem 7.6.4 a heuristically optimal choice (see Remark 7.6.7) for the Euclidean norm. △

3. Filtered evaluation modules and Functional Transcendence

In this section, we develop the structure of the vanishing filtration jumps of the multivariable evaluation modules of auxiliary polynomial functions that we described in § 2.13.1. Their formalism and the basic facts are collected in § 3.1, where we prove the commutativity in the formation of Cartesian products and vanishing filtration jump exponent vectors of the evaluation modules. In § 3.2, we survey some of the literature on the classical Shidlovsky lemma from the historical proof [Shi59] of the Siegel–Shidlovsky theorem on special values of E-functions, and the deeper work of Chudnovsky and Osgood on the functional Schmidt subspace theorem — *Kolchin's problem* — in differential algebra. This finer information simplifies the statements and proofs of our arithmetic holonomy bounds, and it is furthermore indispensable for any refinements to quantitative linear independence measures and Diophantine inequalities. We do nevertheless remark that, following indications in § 7.7, one could in principle dispense with the differential algebra theorems for the particular *qualitative* linear independence proofs in this paper. Finally, just to give a sense of completeness and a proper historical context, we collect in § 3.3 some of the most basic examples of perfect Padé approximants to holonomic functions, which can be considered as a prototype and an introduction to the functional bad approximability theorems collected in § 3.2.

3.1. **The evaluation module for Cartesian products.** We formalize the discussion of § 2.13.1.

3.1.1. *Evaluation module.* Consider a bounded Lebesgue-measurable subset $\Omega \subset [0,\infty)^d$ and a finite indexing set I. In practice we will think of I as a subset of $\{1,\dots,m\}^d$, and so we will use the boldface notation for the index elements $\mathbf{i} \in I$. We fix for the time being an I-tuple of $\mathbf{Q}(\mathbf{x})$-linearly independent formal power series

$$f_{\mathbf{i}}(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!], \qquad \mathbf{x} := (x_1,\dots,x_d), \quad \mathbf{i} \in I.$$

The finite-rank free $\mathbf{Z}$-modules in the following will all depend on the given function $f_{\mathbf{i}}$, which will be considered as fixed and dropped from the notation.

Definition 3.1.2. The *evaluation module* $(E^I_{D,\Omega}, \psi_D)$ defined by the data

$$((f_{\mathbf{i}})_{i\in I}; D; \Omega)$$

is a pair of a finite-rank free $\mathbf{Z}$-module $E^I_{D,\Omega}$ together with a $\mathbf{Z}$-module homomorphism $\psi_D : E^I_{D,\Omega} \to \mathbf{Q}[\![\mathbf{x}]\!]$, constructed as follows. For $E^I_{D,\Omega}$, take the $\mathbf{Z}$-linear span of all I-tuples of monomials $\mathbf{x}^{\mathbf{k}(\mathbf{i})}$ with $\mathbf{k}(\mathbf{i}) \in (D\cdot\Omega)\cap\mathbf{Z}^d$ for all $\mathbf{i}\in I$. For ψ_D, we take the Taylor series development map $\mathbf{Z}$-linearly generated on the basis by the Taylor expansion of $\mathbf{x}^{\mathbf{k}(\mathbf{i})} f_{\mathbf{i}}(\mathbf{x})$:

$$\psi_D \,:\, E^I_{D,\Omega} \otimes_{\mathbf{Z}} \mathbf{Q} \hookrightarrow \mathbf{Q}[\![\mathbf{x}]\!], \qquad \mathbf{x}^{\mathbf{k}(\mathbf{i})} \mapsto \mathbf{x}^{\mathbf{k}(\mathbf{i})} f_{\mathbf{i}}(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!].$$

The evaluation map ψ_D is an *injective* homomorphism, precisely by the $\mathbf{Q}(\mathbf{x})$-linear independence we assumed on the I-tuple $f_{\mathbf{i}}(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!]$.

3.1.3. *Evaluation filtration.* Consider the filtration of the infinite-dimensional $\mathbf{Q}$-vector space $\mathbf{Q}[\![\mathbf{x}]\!]$ of formal power series in d commuting variables $\mathbf{x} := (x_1, \ldots, x_d)$, obtained by firstly grading the monomial basis $\mathbf{x}^{\mathbf{n}}$ by the total degree $|\mathbf{n}| := n_1 + \ldots + n_d$, and then filtering the $\binom{n+d-1}{d-1}$-dimensional $\mathbf{Q}$-vector space homogeneous piece of degree-n elements by the lexicographical ordering of the exponents $\mathbf{n} = (n_1, \ldots, n_d)$ with $|\mathbf{n}| = n$:

$$\mathbf{m} \prec \mathbf{n} \Longleftrightarrow \text{either } |\mathbf{m}| < |\mathbf{n}|, \text{ or } |\mathbf{m}| = |\mathbf{n}| \text{ and } \mathbf{m} \text{ precedes } \mathbf{n} \text{ lexicographically.}$$

We denote the successor function of this total ordering by $\mathbf{n} \mapsto \mathbf{n}^+$. The resulting filtration on

$$\mathbf{Q}[\![\mathbf{x}]\!] =: F = \bigcup_{\mathbf{n} \in \mathbf{N}^d, \prec} F^{(\mathbf{n})}$$

is split and *maximal*: the successor quotient $\mathbf{Q}$-vector spaces $F^{(\mathbf{n})}/F^{(\mathbf{n}^+)} \cong \mathbf{Q}$ are one-dimensional with basis the class of the unique monomial $\mathbf{x}^{\mathbf{n}}$ in $F^{(\mathbf{n})} \smallsetminus F^{(\mathbf{n}^+)}$. Using the monomorphism $\psi_D : E^I_{D,\Omega} \hookrightarrow F$, the $(\mathbf{N}^d, \prec)$-filtration on F induces an $(\mathbf{N}^d, \prec)$-filtration on the $\mathbf{Q}$-vector space $E^I_{D,\Omega} \otimes_{\mathbf{Z}} \mathbf{Q}$:

$$E^{I,(\mathbf{n})}_{D,\Omega} := \psi_D^{-1}\left(F^{(\mathbf{n})}\right) \subset E^I_{D,\Omega} \otimes_{\mathbf{Z}} \mathbf{Q}.$$

The monomorphism $\psi_D : E^I_{D,\Omega} \hookrightarrow F$ induces a *monomorphism* on the graded successive quotients:

$$\psi_D^{(\mathbf{n})} : E^{I,(\mathbf{n})}_{D,\Omega}/E^{I,(\mathbf{n}^+)}_{D,\Omega} \hookrightarrow F^{(\mathbf{n})}/F^{(\mathbf{n}^+)}.$$

and since the codomain of this monomorphism is the one-dimensional $\mathbf{Q}$-vector space $F^{(\mathbf{n})}/F^{(\mathbf{n}^+)}$, it follows that

$$\forall \mathbf{n} \in \mathbf{N}^d, \qquad \dim_{\mathbf{Q}}\left(E^{I,(\mathbf{n})}_{D,\Omega}/E^{I,(\mathbf{n}^+)}_{D,\Omega}\right) \in \{0, 1\}. \tag{3.1.4}$$

The sum of all those $\{0,1\}$ dimensions equals $\dim_{\mathbf{Q}}(E^I_{D,\Omega} \otimes \mathbf{Q}) = \operatorname{rank}(E^I_{D,\Omega})$. It follows that the *vanishing filtration jumps* set

$$\mathcal{V}^I_{D,\Omega} := \left\{\mathbf{n} \in \mathbf{N}^d \,:\, \dim_{\mathbf{Q}}\left(E^{I,(\mathbf{n})}_{D,\Omega}/E^{I,(\mathbf{n}^+)}_{D,\Omega}\right) = 1\right\} \subset \mathbf{N}^d$$

satisfies

$$\#\mathcal{V}^I_{D,\Omega} = \operatorname{rank}(E^I_{D,\Omega}), \quad \text{and} \quad E^{I,(\mathbf{n}^+)}_{D,\Omega} = E^{I,(\mathbf{n})}_{D,\Omega} \text{ for } \mathbf{n} \notin \mathcal{V}^I_{D,\Omega}. \tag{3.1.5}$$

The $\prec$ filtration also shows that $\mathbf{n} \in \mathcal{V}^I_{D,\Omega}$ are exactly the exponents that occur in the monomials $\mathbf{x}^{\mathbf{n}}$ in the leading order jet $|\mathbf{n}| = \operatorname{ord}_{\mathbf{x}=\mathbf{0}}(F)$ of some nonzero element $F \in \psi_D\left(E^I_{D,\Omega})\right) \smallsetminus \{0\}$.

We have proved:

Lemma 3.1.6. *Under the total ordering $(\mathbf{N}^d, \prec)$ and the above premise of the $\mathbf{Q}(\mathbf{x})$-linear independence of the power series $(f_{\mathbf{i}})_{\mathbf{i} \in I}$, there are precisely* $\operatorname{rank}(E^I_{D,\Omega})$ *exponent vectors $\mathbf{n} \in \mathbf{N}^d$ for which there exists a nonzero element $G \in \psi_D\left(E^I_{D,\Omega})\right) \subset \mathbf{Q}[\![\mathbf{x}]\!]$ whose $\prec$-minimal exponent monomial is $c\,\mathbf{x}^{\mathbf{n}}$ for some nonzero $c \in \mathbf{Q}^{\times}$.*

Furthermore, for any $G \in \psi_D\left(E^I_{D,\Omega})\right) \smallsetminus \{0\}$ of $\mathbf{x} = \mathbf{0}$ vanishing order n, and for any nonzero monomial $c\,\mathbf{x}^{\mathbf{n}}$ in G of the minimal degree $|\mathbf{n}| = n$, there exists an $F \in \psi_D\left(E^I_{D,\Omega})\right)$ with $\prec$-minimal monomial $\mathbf{x}^{\mathbf{n}}$.

3.1.7. *Cartesian products.* Consider two bounded Lebesgue-measurable subsets $\Omega_1 \subset [0,\infty)^{d_1}$ and $\Omega_2 \subset [0,\infty)^{d_2}$, a positive integer $D \in \mathbf{N}_{>0}$, two respective index sets I_1 and I_2 as above, and for each $h \in \{1,2\}$, two respective I_h-tuples of $\mathbf{Q}\left(x_1^{(h)},\dots,x_{d_h}^{(h)}\right)$-linearly independent formal power series $\{f_{\mathbf{i}}\}_{\mathbf{i}\in I_1}, \{g_{\mathbf{j}}\}_{\mathbf{j}\in I_2}$. These respective data sets define two evaluation modules $\psi_D^{(h)} : E_{D,\Omega_h}^{I_h} \hookrightarrow \mathbf{Q}[\![\mathbf{x}^{(h)}]\!]$, as well as a (d_1+d_2)-variable evaluation module $E_{D,\Omega_1\times\Omega_2}^{I_1\times I_2}$, the *Cartesian product*, defined by the $\mathbf{Q}(\mathbf{x}^{(1)},\mathbf{x}^{(2)})$-linearly independent $(I_1\times I_2)$-tuple of formal power series $f_{\mathbf{i}}(\mathbf{x})g_{\mathbf{j}}(\mathbf{y})$. There is hence a tautological $\mathbf{Z}$-module isomorphism

$$\begin{aligned} E_{D,\Omega_1}^{I_1} \times E_{D,\Omega_2}^{I_2} &\xrightarrow{\simeq} E_{D,\Omega_1\times\Omega_2}^{I_1\times I_2}, \\ (f_{\mathbf{i}}(\mathbf{x}), g_{\mathbf{j}}(\mathbf{y}))_{\mathbf{i}\in I_1,\mathbf{j}\in I_2} &\mapsto (f_{\mathbf{i}}(\mathbf{x})g_{\mathbf{j}}(\mathbf{y}))_{(\mathbf{i},\mathbf{j})\in I_1\times I_2}, \end{aligned} \tag{3.1.8}$$

under which the evaluation map $\psi_D : E_{D,\Omega_1\times\Omega_2}^{I_1\times I_2} \hookrightarrow \mathbf{Q}[\![\mathbf{x}^{(1)},\mathbf{x}^{(2)}]\!]$ of the product commutes with the product of the evaluation maps

$$(\psi_D^{(1)},\psi_D^{(2)}) : E_{D,\Omega_1}^{I_1} \times E_{D,\Omega_2}^{I_2} \hookrightarrow \mathbf{Q}[\![\mathbf{x}^{(1)},\mathbf{x}^{(2)}]\!].$$

In combination with Lemma 3.1.6, this remarks leads to the following key result, which we can most succinctly express by saying that the formation of Cartesian products of evaluation modules commutes with the formation of their vanishing filtration jumps.

Lemma 3.1.9. *Fix the* $\mathbf{Q}(\mathbf{x}^{(1)})$*-linearly independent* I_1*-tuple* $\left(f_{\mathbf{i}}(\mathbf{x}^{(1)})\right)_{\mathbf{i}\in I_1}$ *and the* $\mathbf{Q}(\mathbf{x}^{(2)})$*-linearly independent* I_2*-tuple* $\left(g_{\mathbf{j}}(\mathbf{x}^{(2)})\right)_{\mathbf{j}\in I_2}$*. Under the notation and premises of the current § 3.1, the vanishing filtration jumps of the associated evaluation modules satisfy*

$$\mathcal{V}_{D,\Omega_1\times\Omega_2}^{I_1\times I_2} = \mathcal{V}_{D,\Omega_1}^{I_1} \times \mathcal{V}_{D,\Omega_2}^{I_2}, \tag{3.1.10}$$

as subsets of $\mathbf{N}^{d_1+d_2} = \mathbf{N}^{d_1}\times\mathbf{N}^{d_2}$.

In view of the importance of this basic lemma for the sequel, we give two proofs.

First proof of Lemma 3.1.9. By (3.1.5) and (3.1.8), the two sets in the asserted equality (3.1.10) are finite and of the same cardinality

$$\operatorname{rank}(E_{D,\Omega_1\times\Omega_2}^{I_2\times I_2}) = \operatorname{rank}(E_{D,\Omega_1}^{I_1})\operatorname{rank}(E_{D,\Omega_2}^{I_2}).$$

It therefore suffices to prove that one of these sets is contained by the other. But

$$\mathcal{V}_{D,\Omega_1}^{I_1} \times \mathcal{V}_{D,\Omega_2}^{I_2} \subseteq \mathcal{V}_{D,\Omega_1\times\Omega_2}^{I_1\times I_2}$$

is made clear by the following product construction. For any pair $\mathbf{n}_1 \in \mathcal{V}_{D,\Omega_1}^{I_1} \subset \mathbf{N}^{d_1}$ and $\mathbf{n}_2 \in \mathcal{V}_{D,\Omega_2}^{I_2} \subset \mathbf{N}^{d_2}$, there exist by definition two auxiliary function evaluations $G_1(\mathbf{x}^{(1)}) \in \psi_D\left(E_{D,\Omega_1}^{I_1}\right) \subset \mathbf{Q}[\![\mathbf{x}^{(1)}]\!]$ and $G_2(\mathbf{x}^{(2)}) \in \psi_D\left(E_{D,\Omega_2}^{I_2}\right) \subset \mathbf{Q}[\![\mathbf{x}^{(2)}]\!]$ such that $\mathbf{n}_1$ is the $\prec$-minimal exponent in $G_1(\mathbf{x}^{(1)}) = c_1\cdot(\mathbf{x}^{(1)})^{\mathbf{n}_1}+\dots$ among the totally ordered exponent set $(\mathbf{N}^{d_1},\prec)$, and $\mathbf{n}_2$ is the $\prec$-minimal exponent in $G_2(\mathbf{x}^{(2)}) = c_2\cdot(\mathbf{x}^{(2)})^{\mathbf{n}_2}+\dots$ among the totally ordered exponent set $(\mathbf{N}^{d_2},\prec)$. Consider then the product

$$G(\mathbf{x}^{(1)},\mathbf{x}^{(2)}) := G_1(\mathbf{x}^{(1)})G_2(\mathbf{x}^{(2)}) \in \mathbf{Q}[\![\mathbf{x}^{(1)},\mathbf{x}^{(2)}]\!].$$

By construction of the Cartesian product, we see that G belongs to the evaluation range $\psi_D\left(E_{D,\Omega_1\times\Omega_2}^{I_2\times I_2}\right)$. It has the nonzero coefficient $c_1c_2 \in \mathbf{Q}^\times$ in the multidegree

$(\mathbf{n}_1,\mathbf{n}_2)\in\mathbf{N}^{d_1}\times\mathbf{N}^{d_2}=\mathbf{N}^{d_1+d_2}$. We claim that for the $(\mathbf{N}^{d_1+d_2},\prec)$ total ordering of the exponents, this is the minimal multidegree in G. It is certainly of the minimal possible vanishing order $|\mathbf{n}_1|+|\mathbf{n}_2|$, for by definition of $\prec$ the factor power series G_1 and G_2 have respective vanishing orders $|\mathbf{n}_1|$ and $|\mathbf{n}_2|$. Now the monomial degrees $(\mathbf{u},\mathbf{v})$ in G having the minimal possible order $|\mathbf{u}|+|\mathbf{v}|=|\mathbf{n}_1|+|\mathbf{n}_2|$ have, by $|\mathbf{u}|\geq|\mathbf{n}_1|$ and $|\mathbf{v}|\geq|\mathbf{n}_2|$, partial degrees $|\mathbf{u}|=|\mathbf{n_1}|$ and $|\mathbf{v}|=|\mathbf{n_2}|$. We have $\mathbf{n}_1\preceq\mathbf{u}$ and $\mathbf{n}_2\preceq\mathbf{v}$. By the definition of the lexicographical ordering, if $(\mathbf{u},\mathbf{v})\neq(\mathbf{n}_1,\mathbf{n}_2)$, it follows that $(\mathbf{n}_1,\mathbf{n}_2)\prec(\mathbf{u},\mathbf{v})$. Hence, through the example of $G(\mathbf{x}^{(1)},\mathbf{x}^{(2)})=c_1c_2\cdot(\mathbf{x}^{(1)})^{\mathbf{n}_1}(\mathbf{x}^{(2)})^{\mathbf{n}_2}+\ldots$, we have found that $(\mathbf{n}_1,\mathbf{n}_2)\in\mathcal{V}^{I_1\times I_2}_{D,\Omega_1\times\Omega_2}$, and in this way we have proved the requisite inclusion $\mathcal{V}^{I_1}_{D,\Omega_1}\times\mathcal{V}^{I_2}_{D,\Omega_2}\subseteq\mathcal{V}^{I_1\times I_2}_{D,\Omega_1\times\Omega_2}$. □

Second proof of Lemma 3.1.9. We can also directly see the reverse inclusion,

$$\mathcal{V}^{I_1\times I_2}_{D,\Omega_1\times\Omega_2}\subseteq\mathcal{V}^{I_1}_{D,\Omega_1}\times\mathcal{V}^{I_2}_{D,\Omega_2}.$$

Let

$$F(\mathbf{x}^{(1)},\mathbf{x}^{(2)})=\sum_{\mathbf{i}\in I_1,\,\mathbf{j}\in I_2}Q_{\mathbf{i},\mathbf{j}}(\mathbf{x}^{(1)},\mathbf{x}^{(2)})f_{\mathbf{i}}(\mathbf{x}^{(1)})g_{\mathbf{j}}(\mathbf{x}^{(2)})\in\psi_D\left(E^{I_1\times I_2}_{D,\Omega_1\times\Omega_2}\right)$$

be an arbitrary auxiliary function evaluation of the Cartesian product module, with $(\mathbf{N}^{d_1+d_2},\prec)$ minimal monomial $\beta\cdot(\mathbf{x}^{(1)})^{\mathbf{n}_1}(\mathbf{x}^{(2)})^{\mathbf{n}_2}$. Then

$$\frac{1}{\mathbf{n}_2!}\frac{\partial^{|\mathbf{n}_2|}}{(\partial\mathbf{x}^{(2)})^{\mathbf{n}_2}}\Big|_{\mathbf{x}^{(2)}=\mathbf{0}}\Big\{F(\mathbf{x}^{(1)},\mathbf{x}^{(2)})\Big\}=\beta\cdot(\mathbf{x}^{(1)})^{\mathbf{n}_1}+\ldots\big[\prec\text{-higher terms}\big],$$

for the monomials of this specialization are exactly the $\gamma\cdot(\mathbf{x}^{(1)})^{\mathbf{k}}$ such that $\gamma\cdot(\mathbf{x}^{(1)})^{\mathbf{k}}(\mathbf{x}^{(2)})^{\mathbf{n}_2}$ are monomials from $F(\mathbf{x}^{(1)},\mathbf{x}^{(2)})$. In this way, $\mathbf{n}_1\in\mathcal{V}^{I_1}_{D,\Omega_1}$. Similarly, $\mathbf{n}_2\in\mathcal{V}^{I_2}_{D,\Omega_2}$, and the requisite inclusion $\mathcal{V}^{I_1\times I_2}_{D,\Omega_1\times\Omega_2}\subseteq\mathcal{V}^{I_1}_{D,\Omega_1}\times\mathcal{V}^{I_2}_{D,\Omega_2}$ is proved. □

We record the main corollary we will use.

Corollary 3.1.11. *Consider a positive integer $D\in\mathbf{N}_{>0}$ and an m-tuple $f_1,\ldots,f_m$ in $\mathbf{C}[\![x]\!]$ of $\mathbf{C}(x)$-linearly independent formal power series. For these data, there exists a sequence*

$$0\leq u(1)<\cdots<u(mD)$$

of mD non-negative integers such that the following holds for every $d=1,2,3,\ldots$:

In every nonzero formal power series of the shape

$$F(\mathbf{x}):=\sum_{\mathbf{i}\in\{1,\ldots,m\}^d}Q_{\mathbf{i}}(\mathbf{x})f_{i_1}(x_1)\cdots f_{i_d}(x_d)\in\mathbf{C}[\![x_1,\ldots,x_d]\!]\smallsetminus\{0\},$$

where $Q_{\mathbf{i}}(x_1,\ldots,x_d)\in\mathbf{C}[x_1,\ldots,x_d]$ are polynomials having all their partial degrees $\deg_{x_j}Q_{\mathbf{i}}<D$, all monomials $\beta\mathbf{x}^{\mathbf{n}}$ with minimal total degree $|\mathbf{n}|=n_1+\ldots+n_d$ have

$$\mathbf{n}\in\big\{0\leq u(1)<\cdots<u(mD)\big\}^d\subset\mathbf{N}^d.$$

Proof. Take the evaluation module $\psi_D:E^{\{1,\ldots,m\}}_{D,[0,1)}\hookrightarrow\mathbf{Q}[\![x]\!]$ defined by the $\mathbf{Q}(x)$-linearly independent power series $f_1,\ldots,f_\in\mathbf{Q}[\![x]\!]$. Clearly, $\operatorname{rank}\left(E^{\{1,\ldots,m\}}_{D,[0,1)}\right)=mD$. We define

$$\big\{0\leq u(1)<\cdots<u(mD)\big\}=\mathcal{V}^{\{1,\ldots,m\}}_{D,[0,1)}\subset\mathbf{N}$$

to be the vanishing filtration jumps for this single-variable evaluation module. By Lemma 3.1.9, it then follows for each $d = 1, 2, 3, \dots$ that the Cartesian d-th power evaluation module $\left(E_{D,[0,1)^d}^{\{1,\dots,m\}^d}, \psi_D\right)$ defined by the $\mathbf{Q}(x_1, \dots, x_d)$-linearly independent formal power series $f_{\mathbf{i}}(\xi) := f_{i_1}(x_1) \cdots f_{i_d}(x_d)$ has vanishing filtration jumps at exactly the d-th Cartesian power set

$$\mathcal{V}_{D,[0,1)^d}^{\{1,\dots,m\}^d} = \left\{0 \le u(1) < \cdots < u(mD)\right\}^d \subset \mathbf{N}^d.$$

The result now follows by Lemma 3.1.6. □

3.2. **Functional bad approximability.** When the functions $f_1, \dots, f_m \in \mathbf{Q}[\![x]\!]$ are holonomic, it turns out possible to almost completely determine the mD vanishing filtration jumps of the ensuing univariate evaluation module $E_D := E_{D,[0,1)}^{\{1,\dots,m\}}$. This is the content of the Chudnovsky–Osgood theorem 3.2.13, which can be seen as a functional analog for holonomic functions of the Roth–Schmidt bad approximability theorem. The roots of all of this are in Hermite's memoir (discussion in § 3.3 below) on the exponential function and the transcendence of the number e.

Basic Remark 3.2.1. For the system $\{f_1, \dots, f_m\} = \{e^{\alpha_1 x}, \dots, e^{\alpha_m x}\}$ of pairwise distinct exponential functions, Hermite [Her1893], in a letter published in 1893 (after having published similar formulas already in [Her1874]), found the explicit $\mathbf{C}[x]$-linear combination of the maximal $x = 0$ vanishing order for an arbitrary degree vector $(D_1, \dots, D_m) \in \mathbf{N}^m$:

$$\begin{aligned}\int_{|z|=R} \frac{e^{xz}\, \mu_{\mathrm{Haar}}(z)}{(z-\alpha_1)^{D_1+1} \cdots (z-\alpha_m)^{D_m+1}} &=: \sum_{i=1}^m P_i(x) e^{\alpha_i x}, \qquad R > \max_i |\alpha_i|, \\ &= \frac{1}{(D_1 + \ldots + D_m + m)!}\, x^{-1+\sum_{i=1}^m (D_i+1)} + O\left(x^{\sum_{i=1}^m (D_i+1)}\right),\end{aligned} \tag{3.2.2}$$

This follows upon unfolding the residue calculus of the complex contour integral and finding the thus-explicitable polynomials $P_1, \dots, P_m$ to have the *exact* degrees $\deg P_i = D_i$. The right-hand side of (3.2.2) follows by D_i partial integrations upon computing the integrand residues at the poles $z = \alpha_1, \dots, \alpha_m$ in the bounded component of $\mathbf{C} \smallsetminus \{|z| = R\}$; on the other hand, the $x = 0$ exact vanishing order development (3.2.2) follows by computing the residue at the unique pole $z = \infty$ in the complementary component. Having for these particular polynomials — the so-called *type I Hermite–Padé approximants* — the exact degrees $\deg P_i = D_i$ (which can furthermore be taken completely arbitrary), and this exact vanishing order (3.2.2), proves by an argument similar to the proof of Lemma 3.1.6 that for *arbitrary* polynomials $Q_1, \dots, Q_m \in \mathbf{C}[x]$, the strongest possible form of (rational) functional bad approximability is in place:

$$\mathrm{ord}_{x=0} \left(Q_1 f_1 + \ldots + Q_m f_m\right) \le \sum_{i=1}^m (\deg Q_i + 1) - 1. \tag{3.2.3}$$

Mahler [Mah68] termed such systems *perfect*, and found a few other examples (incidentally obtainable from Hermite's formula by a substitution and a limit [Chu83b, page 331]), including the binomial system

$$\{(1-x)^{\alpha_1}, \dots, (1-x)^{\alpha_m}\}, \qquad \text{when all } \alpha_i - \alpha_j \notin \mathbf{Z}, \tag{3.2.4}$$

and, under the additional constraint[14] $D_1 \le D_2 \le \cdots \le D_m$, the logarithm system [Mah19b, Jag64]

$$\{1, \log(1-x), \log^2(1-x), \ldots, \log^{m-1}(1-x)\}. \tag{3.2.5}$$

In [Mah53], Mahler used the explicit linear forms for the system (3.2.5) to prove the explicit inequality $|\pi - p/q| > q^{-42}$ for all positive integers $p, q \ge 2$. Gregory Chudnovsky [Chu79, Chu83b, Chu83a] has used (like Thue, Siegel, and Baker before him, cf. § 3.3.3) the systems (3.2.4) and (3.2.5) to derive excellent effective irrationality exponents for suitable roots $\sqrt[n]{b/a}$ from rational numbers, as well as for logarithms of rational numbers. A fairly general class of perfect systems are the *Angelesco–Nikishin systems* [Ang1919, Nik80, Sor96, NS91] in the theory of the Cauchy transform and orthogonal polynomials; their perfection was proven in full generality by Fidalgo Prieto and López-Lagomasino [FPLL11a, FPLL11b]. The fact that the Padé approximants to the polylogarithm system $\{f_1, \ldots, f_m\} = \{1, \mathrm{Li}_1, \mathrm{Li}_2, \ldots, \mathrm{Li}_{m-1}\}$ turn out to be Angelesco–Nikishin systems was at the root[15] of Ball and Rivoal's work [Riv00, BR01] on the arithmetic of zeta values.

In the particular case $D_1 = \cdots = D_m = D - 1$, we can equivalently express the functional bad approximability property (3.2.3) into the framework of § 3.1: it precisely means that the evaluation module E_D has the vanishing filtration jumps set

$$\mathcal{V}_{D,[0,1)}^{\{1,\ldots,m\}} = \{0, 1, \ldots, mD - 1\}. \tag{3.2.6}$$

△

In differential algebra, as we briefly indicated in § 2.13.3, Kolchin [Kol59] proved an analog of Liouville's Diophantine inequality, and asked[16] for an analog of the fundamental theorem on algebraic numbers that Roth had proved four years prior:

Problem 3.2.7 (Kolchin's Problem)**.** *Given a non-rational formal power series solution $f \in \mathbf{C}[\![x]\!] \setminus \mathbf{C}(x)$ of some linear ODE $L(f) = 0$ over $\mathbf{C}(x)$, to prove that $(2+\varepsilon)\max(\deg P, \deg Q) + O_{\varepsilon,f}(1)$ is the highest $x = 0$ vanishing order possible for the error $f(x) - P(x)/Q(x)$ in any rational function approximation.*

In fact, Kolchin's setup was more general and not limited to linear ODEs; he worked in an arbitrary nontrivial valued differential field, and his Liouville inequality [Kol59, § 5] thus also applied to arbitrary (nonlinear) ODEs over $\mathbf{C}(x)$. The first such result, weaker than Kolchin's, appears to be Maillet's [Mai1906, page 266].

In terms of our evaluation modules in § 3.1, Kolchin's Liouville-type result for the case of a formal power series solution to an r^{th} order linear ODE $\mathcal{L}(f) = 0$ can be expressed by saying that the $2D$ vanishing filtration jumps in the module defined by $\{f_1, f_2\} = \{1, f\}$ are contained by $\{0, 1, \ldots, rD + O_{\mathcal{L}}(1)\}$. His (implicit) Roth-type conjecture is that they should in fact be contained by $\{0, 1, \ldots, (2+\varepsilon)D + O_{\varepsilon,\mathcal{L}}(1)\}$ for every $\varepsilon > 0$.

Independently in the same year, Shidlovsky [Shi59] (see also [Shi89, § 3.5, Lemma 8], [Lan66, § VII.3], or [Mah76, § 4]) discovered a more accurate form of the functional

[14] Sometimes termed *weak perfection*.

[15] As a starting or inspiration point, even though they ultimately devised a different (but related) function system, see [FR03, Théorème 1] or [Fis04, § 2.4].

[16] From [Kol59]: "It remains to make the obvious remark, in view of the deep Thue–Siegel–Roth improvement of Liouville's theorem (see K. F. Roth, Mathematika vol. 2 (1955) pp. 1–20), that it would be desirable to obtain a similar improvement in the present theorem."

Liouville inequality in the case of linear ODEs over $\mathbf{C}(x)$, and used it to complete the main results of Siegel's algebraic independence theory [Zan14, Sie49] for special values of E-functions. We reformulate Shidlovsky's lemma into our language of the vanishing filtration jumps.

Theorem 3.2.8 (Shidlovsky)**.** *For the case of a system* $\{f_1, \ldots, f_m\}$ *whose* $\mathbf{Q}(x)$-*linear span is* m-*dimensional and closed under the derivation* d/dx, *there exists a constant* $C = C(f_1, \ldots, f_m)$ *such that, for every* $D \in \mathbf{N}_{>0}$, *the vanishing filtration jumps of the evaluation module* $E_{D,[0,1)}^{\{1,\ldots,m\}}$ *satisfy*

$$\mathcal{V}_{D,[0,1)}^{\{1,\ldots,m\}} \subseteq \{0, 1, \ldots, mD + C\}, \qquad \#\mathcal{V}_{D,[0,1)}^{\{1,\ldots,m\}} = mD. \tag{3.2.9}$$

Although Shidlovsky's original work did not supply an effective procedure to compute the constant C out of the rank-m first-order linear differential system $\mathbf{y}' = A\mathbf{y}$ that has $\mathbf{y} = (f_1, \ldots, f_m)^{\mathrm{t}}$ as a solution, such theorems were eventually obtained, firstly by Chudnovsky [Chu80, Corollary 11.3.10] in the Fuchsian case (which is certainly the case we are concerned with, see Remark 3.2.12 below), and then eventually in the general case by Bertrand, Beukers, Chirskii, and Yebbou, see [BB85] as complemented by [BCY04, § 3]. A far-reaching generalization of Shidlovsky's lemma is in Bertrand [Ber12, Théorème 2].

We state only a crude version of Chudnovsky's result on the Fuchsian case. A brief treatment of this explicit zero estimate is also sketched in André's book [And04, § III, Appendix].

Theorem 3.2.10 (Chudnovsky)**.** *Suppose the system* $\mathbf{f} := \{f_1, \ldots, f_m\} \in \mathbf{Q}[\![x]\!]^m \smallsetminus (x\mathbf{Q}[\![x]\!])^m$ *of* $\mathbf{Q}(x)$-*linearly independent formal power series arises as the full component vector of some solution* $\mathbf{y} = \mathbf{f}^{\mathrm{t}}$ *to a Fuchsian first-order linear differential system* $\mathbf{y}' = A\mathbf{y}$, *where* $A \in M_{m\times m}(\mathbf{Q}(x))$. *Let* $S \subset \mathbf{P}^1$ *be the set of poles in the matrix of rational functions* A, *and define* $h := \sum_{s\in S} \varepsilon_s$, *where* ε_s *is the negative of the smallest real part of any exponent that occurs in the asymptotic development of any one of the functions* $f_i(x)$ *at the regular singular point* $x = s$ *of the Fuchsian ODE.*

Then, for all $D \in \mathbf{N}_{>0}$, *the* mD *vanishing filtration jumps of the evaluation module* $E_{D,[0,1)}^{\{1,\ldots,m\}}$ *are contained by the set*

$$\mathcal{V}_{D,[0,1)}^{\{1,\ldots,m\}} \subseteq \left\{0, 1, 2, \ldots, mD + (\#S - 2)\binom{m}{2} + mh\right\};\ \#\mathcal{V}_{D,[0,1)}^{\{1,\ldots,m\}} = mD. \tag{3.2.11}$$

Remark 3.2.12. Thanks to the work on the global nilpotence property by David and Gregory Chudnovsky [CC85a], [DGS94, § VIII], [And89, § VI], [DV01] and the theorem of Honda and Katz [DGS94, § III.6], [And89, § IV.5.3], the holonomic power series in all our (abstract) theorems in this paper are automatically of the Fuchsian class: they have only regular singular points (with rational exponents). Hence Theorem 3.2.10 applies to them as an explicit Shidlovsky bound. For the proofs of Theorem 2.5.1 and all its generalizations, this theorem is already sufficient under the supplemental assumption — which is satisfied in all the applications we could conceive of — that the $\mathbf{Q}(x)$-linear span of $f_1, \ldots, f_m$ is closed under the derivation d/dx. This remark could also have a significance for the project of refining our qualitative linear independence results to quantitative measures of linear independence.

At the same time, the proof of the Chudnovskys's theorem (namely: of Galočkin's canceling factorials property and the global nilpotence of an integrable connection that admits at least one G-series formal solution with $\mathbf{C}(x)$-linearly independent components) *itself* relies on a suitable qualitative Shidlovsky lemma [DGS94, Prop. VIII.2.3], [CC85a, Theorem 3.1, Lemma 8.3], [And89, § VI.2] in the dual form for simultaneous — that is now *type II Hermite–Padé* — functional rational approximants $f_i \approx P_i/Q$, $1 \leq i \leq m$, selected to have integer coefficients of controlled size as provided by the Thue–Siegel lemma. △

All these theorems also embed as very special cases into the broader subject of zero multiplicity estimates for functions satisfying a possibly nonlinear algebraic differential system. This path was opened up by the groundbreaking works of Nesterenko [Nes88] and Brownawell–Masser [BM80]. We refer to Binyamini [Bin16] for a survey, a modern treatment, and refinements of a large portion of the literature on this rather vast topic; and to [Nes96, NP01] for applications to algebraic independence. In the special context of linear ODEs, the separate streams opened up by Kolchin and Shidlovsky converged in the early 1980s with the resolution of Problem 3.2.7, independently by David and Gregory Chudnovsky [CC83] and Osgood [Osg85].

Theorem 3.2.13 (Chudnovsky, Osgood)**.** *Consider an arbitrary set $\{f_1,\dots,f_m\}$ of holonomic functions in $\mathbf{Q}[\![x]\!]$. That is, our only assumption now is that each of the formal power series $f_i(x)$ separately satisfies some nonzero linear ODE $\mathcal{L}_i(f_i)=0$. For an arbitrary $\varepsilon>0$, there exists a constant $C(\varepsilon)=C(\varepsilon;\mathcal{L}_1,\dots,\mathcal{L}_m)$, effectively computable from the arguments in* [CC83, § 2], *such that for all $D\in\mathbf{N}_{>0}$, the vanishing filtration jumps of the evaluation module $E_D = E^{\{1,\dots,m\}}_{D,[0,1)}$ satisfy*

$$\mathcal{V}^{\{1,\dots,m\}}_{D,[0,1)} \subseteq \big\{0,1,2,\dots,(m+\varepsilon)D+C(\varepsilon)\big\}, \qquad \#\mathcal{V}^{\{1,\dots,m\}}_{D,[0,1)} = mD.$$

In conjunction with Corollary 3.1.11, these theorems can be summarized into the following proposition.

Lemma 3.2.14. *Let $f_1,\dots,f_m\in\mathbf{C}[\![x]\!]$ be $\mathbf{C}(x)$-linearly independent* holonomic *power series: there exist nonzero linear differential operators $\mathcal{L}_i$ over $\mathbf{Q}(x)$ with $\mathcal{L}_i(f_i)=0$. Then, for every $\varepsilon>0$, there exists a constant $C(\varepsilon)\in\mathbf{R}$, in principle effectively computable from the datum $(\varepsilon;\mathcal{L}_1,\dots,\mathcal{L}_m)$ alone, such that the following is true.*

We consider an arbitrary positive integer $d\in\mathbf{N}_{>0}$, and write

$$\mathbf{x} := (x_1,\dots,x_d), \qquad f_{\mathbf{i}}(\mathbf{x}) := \prod_{s=1}^{d} f_{i_s}(x_s) \quad \text{for } \mathbf{i} := (i_1,\dots,i_d)\in\{1,\dots,m\}^d.$$

Consider further an arbitrary positive integer $D\in\mathbf{N}_{>0}$ and, over $\mathbf{i}\in\{1,\dots,m\}^d$, an arbitrary set of polynomials

$$Q_{\mathbf{i}}(\mathbf{x})\in\mathbf{C}[x_1,\dots,x_d] \;\text{ with }\; \deg_{x_j} Q_{\mathbf{i}} < D \text{ for all } j\in\{1,\dots,d\} \text{ and } \mathbf{i}\in\{1,\dots,m\}^d.$$

Then, in the nonzero formal power series

$$F(\mathbf{x}) := \sum_{\mathbf{i}\in\{1,\dots,m\}^d} Q_{\mathbf{i}}(\mathbf{x}) f_{\mathbf{i}}(\mathbf{x}) \in \mathbf{C}[\![x_1,\dots,x_d]\!]\smallsetminus\{0\},$$

every lowest-order nonzero monomial term $\beta\,\mathbf{x}^{\mathbf{n}}$ in $F(\mathbf{x})$ has necessarily an exponent vector $\mathbf{n} = (n_1, \ldots, n_d)$, all of whose components satisfy

$$n_j \leq (m + \varepsilon)D + C(\varepsilon).$$

If moreover the $\mathbf{Q}(x)$-linear span of $f_1, \ldots, f_m$ is closed under the derivation d/dx, then $\varepsilon = 0$ could be taken.

Remark 3.2.15. David and Gregory Chudnovsky conjecture [CC83, page 5161] that $\varepsilon = 0$ could be taken in Theorem 3.2.13, and therefore — as a consequence — also in Lemma 3.2.14. However, this conjecture remains unproved even for the case [Wan04] of algebraic functions. △

At this point, for the logic of the proofs, the reader may skip directly ahead to § 4. The remainder of § 3 collects some examples, placed in their historical context, behind the theorems that we borrowed without proof in § 3.2.

3.3. **Some explicit constructions of Hermite–Padé approximants.** For the rest of § 3, we collect a few simplest and most fundamental illustrating examples, aiming at a modest attempt at sketching the historical seeds of some of the basic ideas in the proofs of the theorems on functional bad approximability that we collected in § 3.2, but also of the broader concept of holonomy bounds and the way we use them in our present paper. A quintessential illustrating example for the key point in the proofs of Shidlovsky type theorems on functional bad approximability can be taken as the explicit (in the simple case outlined here) determinantal identity (3.3.6) from the theory of the hypergeometric ODE. The number-theoretic relevance of such identities was found by Thue when he created the subject of non-effective Diophantine approximation. Our approach here to Apéry limits has perhaps some faint similarity to Thue's paradigm with its organic ineffectivity; the proofs that we have of the explicit holonomy bounds of Theorem 2.5.1 do not[17] contain, even in principle, any effective procedure for the far more elusive problem of outputting a set of $\mathbf{Q}(x)$-vector space generators for the finite-dimensional holonomic module $\mathcal{H}(b_1, \ldots, b_r; \varphi)$ attached to a given holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ paired up to a given denominators type $\prod_{i=1}^{r}[1, \ldots, b_i n]$ subject to $|\varphi'(0)| > e^{b_1 + \cdots + b_r}$. Yet, when favored by the presence of suitable anchors (such as we have in § 10) and levers (such as we have in § 9, § 12.1 and § 14.2), the Diophantine repellency principles can occasionally be turned around into true Diophantine inequalities and linear independence proofs. With Thue's method, it took over seventy years until a fairly general-scope theory, on a scale comparable to the Gelfond–Baker method of linear forms in logarithms, started to emerge at the hands of Bombieri and his coauthors [Bom82, BM83, Bom93, BC97a].

But we should probably begin this discussion by delving a bit into our subject's proper origin: the work of Hermite by which he proved the transcendence of e.

[17] Except in the case $\mathbf{b} = \mathbf{0}$ of integer coefficients. In that very special case, even a much more precise integral finiteness counterpart is contained in the work of Bost and Charles [BC22, § 9.1], in an implicitly effective form.

3.3.1. *Hermite approximations.* The memoir [Her1874] on the exponential function was based on the explicit Hermite–Padé approximants to the functions 1, e^x, e^{2x}, …, e^{rx} in order to prove the transcendence of e by specializing $x := 1$. For $r = 1$ the formula is

$$\begin{aligned}
&{}_1F_1\left[\begin{matrix}-m\\ -m-n\end{matrix};x\right] - e^x\cdot {}_1F_1\left[\begin{matrix}-n\\ -m-n\end{matrix};-x\right]\\
&= -\frac{e^x\int_0^1 e^{-tx}t^m(t-1)^n\,dt}{(m+n)!}\,x^{m+n+1}\\
&= (-1)^{n-1}\frac{m!n!}{(m+n)!(m+n+1)!}\,x^{m+n+1} + O(x^{m+n+2})
\end{aligned} \tag{3.3.2}$$

for the unique (up to scalar multiple) combination $B(x) - e^x A(x)$ with $\deg A \leq n, \deg B \leq m$ that vanishes at $x = 0$ to order at least $m + n + 1$. As we can see from the explicit formula, the vanishing order is in fact exactly equal to $m+n+1$. The existence of such a regular array of formulas further proves that for *any* pair $A(x), B(x)$ of nonzero polynomials of degrees $n = \deg A$ and $m = \deg B$, the combination $B(x) - e^x A(x)$ has $x = 0$ vanishing order at most $m+n+1$, with equality if and only if the form $B(x) - e^x A(x)$ is a scalar multiple of (3.3.2). This means that the holonomic function e^x is very badly approximable by rational functions. Hermite's philosophy, which was later taken up by Siegel who started his 1929 paper [Zan14] in exactly the same way as Hermite [Her1874] — outlining an analogy between numbers, to be approximated in the archimedean absolute value, and functions, to be expanded in power series and approximated in terms of the $x = 0$ vanishing order, — was that the functional formulas could be specialized at algebraic arguments to yield a full set of small linear forms with integer coefficients in the numbers (the special values) of interest; which in turn can often be used to prove the $\mathbf{Q}$-linear independence of those numbers. The bad approximability property serves as the sieve for expressing and recognizing a full (linearly independent) set of linear forms, both in the holonomic functions and in their special values, once these are constructed to be reasonably small: as in the functional formula (3.3.2) and its specializations at the algebraic arguments. In Hermite's method, the functional bad approximability of e^x (suitably generalized to include all the powers $1, e^x, e^{2x}, \ldots, e^{rx}$), *via* identities such as (3.3.2), can be specialized at $x := \alpha \in \overline{\mathbf{Q}}^\times$ to derive the Hermite–Lindemann–Weierstrass theorem on the transcendence, and furthermore the bad approximability, of the special value e^α.

The content of Shidlovsky's lemma 3.2.8 and the Chudnovsky–Osgood theorem 3.2.13 can approximately be described as the statement that a property (only very slightly relaxed) of bad approximability by rational functions is in place for any set of holonomic functions. We illustrate this on the most classical cases of perfect systems.

3.3.3. *The Hermite–Padé approximants to* $(1-x)^\nu$. We have the hypergeometric polynomials identity of Jacobi [Jac1859, § 8] to describe explicitly the Padé table for the binomial function (cf. [Zan14, page 75], or Siegel's introductory paper for

Thue's Selected Works volume [Thu77, § 2]):

$$\begin{aligned}
&{}_2F_1\left[\begin{matrix}-\nu-n \quad -m\\ -m-n\end{matrix};x\right]-(1-x)^{\nu}\cdot {}_2F_1\left[\begin{matrix}\nu-m \quad -n\\ -m-n\end{matrix};x\right]\\
&=(-1)^n\frac{\binom{m+\nu}{m+n+1}}{\binom{m+n}{n}}\,{}_2F_1\left[\begin{matrix}-\nu+n+1 \;\; m+1\\ m+n+2\end{matrix};x\right]x^{m+n+1} \qquad (3.3.4)\\
&=(-1)^n\frac{\binom{m+\nu}{m+n+1}}{\binom{m+n}{n}}\,x^{m+n+1}+O(x^{m+n+2}),
\end{aligned}$$

proved for instance by verifying that all three terms satisfy the second-order Gauss hypergeometric equation with parameters $\alpha=-\nu-n, \beta=-m, \gamma=-m-n$, and are therefore $\mathbf{C}$-linearly dependent (the argument is also in [Sie37, Hilfssatz 1]). In this identity, the hypergeometric series on the left-hand side (3.3.4) terminate to polynomials of degrees m and n, and so they give precisely the $[m/n]$ Hermite–Padé approximant $B_{m,n}(x)-(1-x)^{\nu}A_{m,n}(x)$ to the binomial function $(1-x)^{\nu}$, for any $\nu\in\mathbf{C}\smallsetminus\mathbf{Z}$. As the x^{m+n+1} coefficient in the formula (3.3.4) is nonzero, we see here another explicit example of a bad approximability by rational functions.

Now with $\nu\in\mathbf{Q}$, a standard game of Diophantine approximation, both in ineffective (the original and simplest proof of Thue's theorem for the special case of r-th roots from rational numbers) and, in favorable rare circumstances, effective works (Thue, Siegel, Baker, and Gregory Chudnovsky [Chu83b]), is to take the diagonal $[n/n]$ of the Padé table, specialize x to some rational number $\xi\in(0,1)\cap\mathbf{Q}$, and exploit the ensuing small linear forms whose generating function

$$\begin{aligned}
&\sum_{n=0}^{\infty}\big(B_{n,n}(\xi)-(1-\xi)^{\nu}A_{n,n}(\xi)\big)\,z^n\\
&=\sum_{n=0}^{\infty}\left({}_2F_1\left[\begin{matrix}-\nu-n \quad -n\\ -2n\end{matrix};\xi\right]-(1-\xi)^{\nu}{}_2F_1\left[\begin{matrix}\nu-n \quad -n\\ -2n\end{matrix};\xi\right]\right)z^n \qquad (3.3.5)\\
&\in\bigoplus_{n=0}^{\infty}\frac{z^n}{(\mathrm{den}(\nu)\mathrm{den}(\xi))^{2n}\binom{2n}{n}}\,\mathbf{Z}+(1-\xi)^{\nu}\bigoplus_{n=0}^{\infty}\frac{z^n}{(\mathrm{den}(\nu)\mathrm{den}(\xi))^{2n}\binom{2n}{n}}\,\mathbf{Z}
\end{aligned}$$

is holonomic on $\mathbf{C}\smallsetminus\left\{\left(\frac{1\pm\sqrt{1-\xi}}{2}\right)^{-2}\right\}$ and overconvergent at the smaller of these singularities $\left(\frac{1+\sqrt{1-\xi}}{2}\right)^{-2}$; so the convergence disc of (3.3.5) is $|z|<\left(\frac{1-\sqrt{1-\xi}}{2}\right)^{-2}$, the distance to the next singularity. Baker [Bak64] famously used the $\{2,3,\infty\}$-adic properties of (3.3.5) with the choice $\nu:=-1/3$ and $\xi:=3/128$ — so $(1-\xi)^{\nu}=(8/5)\sqrt[3]{2}$, and the 3-adic convergence radius is $1/\sqrt{3}$, thanks to $|\xi|_3=|3/128|_3=1/3$, rather than the "generic" $1/3\sqrt{3}$, — to derive the explicit sub-Liouville inequality $|\sqrt[3]{2}-p/q|>10^{-6}q^{-2.995}$. (Baker's analysis is synthesized by [Chu83b, Theorem 3.5], following which Chudnovsky sets forth to compute the exact denominators asymptotic to refine the crude 4^n from $\binom{2n}{n}$, and thus improve Baker's effective irrationality measure to 2.43; see also [Chu79].)

For the general cubic (or higher) root $\sqrt[r]{a/b}$, this analysis stands no chance for a sub-Liouville effective irrationality measure (unless b is much bigger than a). But Thue [Thu77, § 9], in his groundbreaking paper *Bemerkungen über gewisse Näherungsbrüche algebraischer Zahlen* written in 1907, proved the ineffective irrationality measure $1+r/2+\epsilon$ by — in effect — observing that one excellent rational approximant $p/q \approx \sqrt[r]{a/b} \in \mathbf{Q}^\times \cap (0,1)$ yields the infinite set of fair rational approximants

$$\frac{pB_{n,n}\,(1-aq^r/bp^r)}{qA_{n,n}\,(1-aq^r/bp^r)} \approx \sqrt[r]{\frac{a}{b}},$$

and that these form a fairly dense net of fair approximants, thus precluding — by the gap principle — the existence of a second excellent $p'/q' \approx \sqrt[r]{a/b}$. For that Thue used the $x := 1 - aq^r/bp^r$ specialization in the polynomial identity (cf. [Sie37, Hilfssatz 2], or [AR80, Lemma 2] for an axiomatization)

$$A_{n,n}(x)B_{n+1,n+1}(x) - A_{n+1,n+1}(x)B_{n,n}(x) = (-1)^{n-1}\frac{(n!)^2}{(2n)!(2n+1)!}\,x^{2n+1}, \tag{3.3.6}$$

with the nonvanishing determinant proving at once the requisite non-equality

$$\frac{pB_{n+1,n+1}\,(1-aq^r/bp^r)}{qA_{n+1,n+1}\,(1-aq^r/bp^r)} \neq \frac{pB_{n,n}\,(1-aq^r/bp^r)}{qA_{n,n}\,(1-aq^r/bp^r)}, \qquad \text{for all } n = 0,1,2,\ldots.$$

For arbitrary algebraic targets $\alpha \in \overline{\mathbf{Q}}$ (other than r-th roots or cubic irrationalities), where Thue could not rely on the explicit Hermite–Padé approximants to the binomial functions $(1-x)^\nu$, he instead employed [Thu77, § 11] the Dirichlet box principle, in a flash of insight in the 1908 paper *Om en generel i store hele tal uløsbar ligning*, to derive the existence of similar (but vaguer) polynomial identities. His key discovery was that the inexplicit polynomial identities found nonconstructively by the Dirichlet box principle worked, *grosso modo*, in essentially the same way as in the explicit special case of (3.3.4) and (3.3.6). In particular, Thue used a Wronskian determinant to replace the explicit determinant (3.3.6), now evaluating to some nonzero degree-$2n+1$ polynomial of the form $x^{(2-\eta)n}V(x)$, with small coefficients, for a suitably small parameter $\eta > 0$. As $\deg V \leq \eta n + 1$ — or alternatively, as Thue argued, since the coefficients of V are small, — the polynomial $V(x)$ has forcibly a low order of vanishing at the point $x = 1 - aq^r/bp^r$. Then Thue runs the construction after taking the corresponding derivative of his auxiliary polynomials. (See also Zannier [Zan09, § 2] or Masser [Mas16, § 12] for a detailed treatment and a discussion of nuances.)

Shidlovsky's lemma is a different generalization of Thue's Wronskian argument, whose proofs can still roughly be summarized by a (higher rank) determinantal identity akin to (3.3.6) (of which the latter is strictly speaking a particular and representative case) remaining "almost in the monomial form." It includes Theorems 3.2.8 and 3.2.10, and their multiple variations such as [Bom81, § 3] from the proof of Bombieri's G-functions theorem that we discuss in § 15.1, and [DGS94, Prop. VIII.2.3], [CC85a, Theorem 3.1, Lemma 8.3] from the proof of the Chudnovskys's fundamental theorem that we mentioned in Remark 3.2.12.

3.3.7. *The Hermite–Padé approximants to* $\log(1-x)$. We have [Jac1859, § 8] (see also Feldman–Nesterenko [PS98, ch. 2, § 3.2], Jager [Jag64], and Chu [Chu05] for various generalizations)

$$\begin{aligned} &2\sum_{k=0}^{n}\binom{n}{k}^2 (H_{n-k} - H_k)\,(1-x)^k + \log(1-x)\sum_{k=0}^{n}\binom{n}{k}^2 (1-x)^k \\ &= x^{2n+1}\int_0^1 \frac{t^n(t-1)^n}{(tx-1)^{n+1}}\,dt = -\frac{x^{2n+1}}{(2n+1)\binom{2n}{n}} + O(x^{2n+2}), \end{aligned} \tag{3.3.8}$$

where $H_r := \sum_{k=1}^r 1/k$ are the harmonic numbers.

Remark 3.3.9. In terms of the general Meijer G function

$$G^{m,n}_{p,q}\left(\begin{smallmatrix} a_1 & \cdots & a_p \\ b_1 & \cdots & b_q \end{smallmatrix}\,\middle|\, z\right) = \int_{\Re(s)=\sigma} \frac{\prod_{j=1}^m \Gamma(b_j - s)\prod_{j=1}^n \Gamma(1-a_j+s)}{\prod_{j=m+1}^q \Gamma(1-b_j+s)\prod_{j=n+1}^p \Gamma(a_j - s)}\, z^s \,\frac{ds}{2\pi i},$$

the remainder term in (3.3.8) can be expressed also as $-(n!)^2\, G^{2,0}_{2,2}\left(\begin{smallmatrix} n+1 & n+1 \\ 0 & 0 \end{smallmatrix}\,\middle|\, 1-x\right)$. This general definition as a Barnes integral is valid under the assumption that all poles of all $\Gamma(b_j - s)$ are on the right of the integration line $\Re(s) = \sigma$, while all poles of all $\Gamma(1-a_j+s)$ are on the left of that line. △

Here,

$$\begin{gathered} \sum_{k=0}^n \binom{n}{k}^2 (1-x)^k = {}_2F_1\left[\begin{matrix} -n & -n \\ & 1 \end{matrix}; 1-x\right] = x^n P_n\left(\frac{2-x}{x}\right), \\ P_n(x) := \frac{1}{2^n\, n!}\left(\frac{d}{dx}\right)^n (x^2-1)^n \end{gathered} \tag{3.3.10}$$

in terms of the Legendre polynomials $P_n(x)$: the complete orthogonal system on $[-1,1]$ under the Lebesgue measure and the normalization $P_n(1) = 1$. Their generating series

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^\infty P_n(x) z^n \tag{3.3.11}$$

is precisely the function whose integrality properties — namely: that P_n is integer-valued on the odd integers, amounting to the $\mathbf{Z}[x]$ polynomials $(2x)^n P_n(1/x)$ in § 3.3.13 below — we exploit in § 14.

If like in § 3.3.3 we multiply (3.3.8) by z^n and sum the generating series over $n \in \mathbf{N}$, the resulting

$$\mathbf{Q}[x][\![z]\!] + \log(1-x)\,\mathbf{Z}[x][\![z]\!]$$

function is holonomic in z and has its $\mathbf{Z}[x][\![z]\!]$ and $\mathbf{Q}[x][\![z]\!]$ components satisfy the homogeneous and inhomogeneous first-order ODEs

$$\begin{gathered} (-1+4z-2xz-x^2z^2)Y'(z) + (2-x-x^2z)Y(z) = 0 \\ \text{and} \\ (-1+4z-2xz-x^2z^2)Y'(z) + (2-x-x^2z)Y(z) = -x, \end{gathered} \tag{3.3.12}$$

respectively. These are holonomic functions on

$$\mathbf{C} \setminus \{p_-(x), p_+(x)\}, \qquad p_\pm(x) := \left(\frac{1 \pm \sqrt{1-x}}{x}\right)^2,$$

where these singularities can be also directly obtained from (3.3.10) and (3.3.11).

Specializing $x = 1/n$ and $y = 1/m$, we have

$$p_-(1/n)p_+(1/m)/|mn| = 1 + o_{|m/n| \to 1}(1).$$

This asymptotic is related to the analyticity mechanism with Hadamard products in § 14.1, and could be also used there as an alternative, but ultimately equivalent given § 3.3.13 just below, proof of Theorem C.

3.3.13. *The Hermite–Padé approximants to* $\log\left(\frac{1-x}{1+x}\right)$. The change of variables $x \mapsto 2x/(1+x)$ in (3.3.8) rewrites the formula thus:

$$
\begin{aligned}
&2\sum_{k=0}^{n}\binom{n}{k}^2 (H_{n-k}-H_k)\,(1-x)^k(1+x)^{n-k} + \log\left(\frac{1-x}{1+x}\right)(2x)^n P_n(1/x) \\
&= (x+x^2)^{2n+1}\int_0^1 \frac{t^n(t-1)^n}{(2tx-1-x)^{n+1}}\,dt = -\frac{x^{2n+1}}{(2n+1)\binom{2n}{n}} + O(x^{2n+2}).
\end{aligned}
$$

In § 14 we use the generating functions of these formulas specialized to $x := 1/a$ with a a large odd integer.

4. Concentration of measure

If we randomly and independently sample a very large number $n-1$ of uniformly distributed points of the segment $[0,1]$, the n spacings that remain will be almost surely close to some ordering of the set $\{\log(n/j)/n : 1 \leq j \leq n\}$, while the $n-1$ sample points themselves will be almost surely close to some ordering of the set $\{j/n : 1 \leq j < n\}$. These facts are simplest expressions of the Law of Large Numbers in statistics, with the precise quantitative decay rates being captured by the *concentration of measure phenomenon* of Dvoretzky and Milman [MS86, Mil92, Led01] for the high-dimensional ℓ^r-ball, in the respective cases $r=1$ and $r=\infty$. A popular expression of the measure concentration principle, due to Gromov [Gro07, § $3\frac{1}{2}$.20], is to say that the *observable diameter* of the unit volume ℓ^r-ball in the asymptotic of high dimension n is on the order of only $\frac{1}{\sqrt{n}} = o(1)$, in contrast to its diameter as a metric space which is on the order of $\sqrt{n}$. It is the observable and not the metric properties that are relevant to the various auxiliary polynomial constructions undertaken in Diophantine approximation.

These specific distributions (and the finer statistics) are best expressed by the fact [BGMN05, Theorem 1] (going back[18] to Émile Borel [Bor1914, § V] for the $r=2$ case of the Euclidean ball; see also [SZ90, Lemma 1] or [RR91, § 3]) that the normalized volume measure of the n-dimensional ℓ^r ball is generated stochastically by the random vector

$$
\left(\frac{X_1}{(|X_1|^r+\ldots+|X_n|^r+Z)^{1/r}},\ldots,\frac{X_n}{(|X_1|^r+\ldots+|X_n|^r+Z)^{1/r}}\right),
$$

where $X_1,\ldots,X_n$ are independent and identically distributed random variables with probability density function $\frac{1}{2\Gamma(1+1/r)}e^{-|t|^r}$, and Z is a jointly independent random variable with the exponential density function $e^{-t}\cdot\chi_{[0,\infty)}(t)$. Moreover, the concentration function is Gaussian. These features are general, while for our purposes here, only the simplest statement with the ℓ^∞-ball is used. In this $r\to\infty$ limiting case, one additional (but only technical) simplification is that the random vector components in the stochastic generation of $\mu_{[-1,1]^n}$ are independent rather than merely asymptotically independent.

In the multivariable auxiliary Diophantine constructions §§ 6 and 8, we will use measure concentration ideas as described in § 2.13.4. One aspect of this is

[18] The classic theorem relating the normal distribution to the Euclidean sphere is popularly ascribed to Poincaré in 1912, but see [DF87, § 6] for a scrupulous historical research, and a discussion of a broader context.

to constrict the component sets $\{k_j\}$ of the high-dimensional exponent vectors $\mathbf{k}$ in all the monomials $\mathbf{x}^{\mathbf{k}}$ occurring in the auxiliary polynomial constructions from the evaluation modules that we introduced in § 2.13.1 and studied in § 3.1. This type of application is among the most standard in Diophantine approximation, after the classic works of Roth [BG06, § 6.3.5] and Schmidt[19] [BG06, § 7.5.15], and especially Wirsing [Wir71, § 4.2] (see also [Sto74, Theorem 7.2.1] for two alternative and more detailed treatments of the relevant material from Wirsing's argument). Methodologically our high-dimensional Diophantine analysis in § 6 is rather similar to the multivariable auxiliary polynomial constructions that Wirsing used in the proof of his theorem on the bad approximability of a fixed algebraic target by algebraic approximants of a given degree.

4.1. **The Erdös–Turán bound.** Recall the definition of the box discrepancy function on the hypercube; cf [CDT21, § 2.5.3].

Definition 4.1.1. The (normalized, box) **discrepancy function** $D : [0,1]^n \to (0,1]$ is the supremum over all closed intervals $I = [a,b] \subset [0,1]$ of the defect between the length $\mu_{\mathrm{Lebesgue}}(I) = b - a$ of I and the proportion of points falling inside I:

$$D(t_1,\ldots,t_n) := \sup_{I\subset[0,1)} \left|\mu_{\mathrm{Lebesgue}}(I) - \frac{1}{n}\#\{i \,:\, t_i \in I\}\right|.$$

With the identification $[0,1)^n \xrightarrow{\simeq} \mathbf{T}^n$ induced from $e(t) := \exp(2\pi i t)$, harmonic analysis on the circle supplies a basic way to upper-bound the discrepancy function. The Erdös–Turán inequality states [DT97, Theorem 1.14]

$$D(t_1,\ldots,t_n) \le 3\left(\frac{1}{K+1} + \sum_{k=1}^{K}\frac{1}{k}\left|\frac{e(kt_1)+\ldots+e(kt_n)}{n}\right|\right) \quad \forall K \in \mathbf{N}, \tag{4.1.2}$$

in terms of the character sums on the group $\mathbf{T}$.

4.2. **The large deviations bound.** The following estimate will be critical.

Theorem 4.2.1. *There exist two absolute constants $C, c \in \mathbf{R}$ such that for any $\varepsilon > 0$ and any $n \in \mathbf{N}$, the set*

$$B^n_\varepsilon := \{\mathbf{t} \in [0,1]^n \,:\, D(\mathbf{t}) \ge \varepsilon\}$$

has n-dimensional Lebesgue measure smaller than $Ce^{-c\varepsilon^4 n}$.

For instance, the proof will show that we can take $c = 1/300$ and $C = 100$ in this theorem.

---

[19] In the proof of Schmidt's Subspace theorem, the "$d+1$" exponent for the bad approximability in projective space $\mathbf{P}^d$ receives a probabilistic explanation as the reciprocal of the equal expectations of the individual coordinates of a point ξ taken at random from the surface boundary of the $d+1$-dimensional standard simplex. The concentration property, used for the parameter count at the auxiliary polynomial construction in the Thue–Siegel lemma, states precisely that all the column sums in a tall $n \times (d+1)$ matrix made of $n \to \infty$ such independent and identically distributed random rows ξ converge in probability to the expectation $n/(d+1)$, at an asymptotic rate exponential in $-n$. See [FW94, § 3, *Example 1*] for a broader context of Harder–Narasimhan filtrations on graded algebras of auxiliary functions. The work of Faltings and Wüstholz made a deeper use of probability measures which, in combination with the Faltings product theorem for directly treating the nonvanishing of the auxiliary construction at the special point, ultimately eliminates the difficult geometry of numbers part from Schmidt's proof.

Remark 4.2.2. The existence of an exponential (in the negative of the dimension) asymptotic decay rate is a hallmark of basic concepts of entropy in the theory of large deviations [Ell06]. The specific rate estimate worked out in Theorem 4.2.1 is of no consequence for our purposes, but its existence is used crucially in §§ 6, 8. An alternative path to Theorem 4.2.1, not using harmonic analysis and the Erdös–Turán bound (and with different, indeed better numerical constants c, C), but instead taking for base the rudimentary Chebyshev estimate [Wir71, Lemma 12], can be derived from the bound $\leq e^{-\pi r^2}$ on the concentration function [Led01, Prop. 2.8] for the uniform measure on $[0,1]^n$. In Ledoux's book, the latter concentration inequality is obtained as a consequence [Led01, Cor. 2.6] under a contraction of the sharp estimate $\leq e^{-r^2/2}$ for the concentration function of the canonical Gaussian measure on $\mathbf{R}^n$. The latter, in turn, is traditionally a consequence of Lévy's isoperimetric inequality on the Euclidean sphere [Led01, Theorem 2.3]. △

Our proof of Theorem 4.2.1 will be based on the most standard form of Hoeffding's concentration inequality [Hoe63]. For the sum of independent random variables $X_1, \dots, X_n$ taking values in the interval $[-1,1]$, Hoeffding's inequality [BLM13, Theorem 2.8] bounds the large deviation tail probability exponentially by

$$\mathbf{P}\left(\Big|X_1+\ldots+X_n-\mathbf{E}[X_1+\ldots+X_n]\Big| \geq \varepsilon n\right) \leq 2e^{-\varepsilon^2 n/2}. \tag{4.2.3}$$

On changing ε to $\varepsilon/2$ and using the triangle inequality and the subadditivity of probability, we can apply this to the real and imaginary parts of $\mathbf{T}$-valued independent random variables $Z_1, \dots, Z_n$ to get the following variant:

Lemma 4.2.4 (Hoeffding)**.** *The sum of independent random variables $Z_1, \dots, Z_n$ taking values in the complex unit circle $\mathbf{T}$ has the tail probability large deviations bound*

$$\mathbf{P}\left(\Big|Z_1+\ldots+Z_n-\mathbf{E}[Z_1+\ldots+Z_n]\Big| \geq \varepsilon n\right) \leq 4e^{-\varepsilon^2 n/8}. \tag{4.2.5}$$

Proof of Theorem 4.2.1. In combination with the Erdös–Turán bound, we derive a proof of the theorem, with the following explicit estimate. Take $Z_1, \dots, Z_n$ to be independent and uniformly distributed points of the circle $\mathbf{T}$. Then $\mathbf{E}[Z_1^k+\ldots+Z_n^k]=0$ for all $k=1,2,\ldots$, giving uniformly by Hoeffding's bound (4.2.5)

$$\mathbf{P}\left(\frac{1}{k}\left|\frac{Z_1^k+\ldots+Z_n^k}{n}\right| \geq \varepsilon\right) \leq 4e^{-\varepsilon^2 k^2 n/8}. \tag{4.2.6}$$

It follows that for every $K \in \mathbf{N}$ and $\varepsilon > 0$ the probability

$$\mathbf{P}\left(\frac{3}{K+1}+3\sum_{k=1}^{K}\frac{1}{k}\left|\frac{Z_1^k+\ldots+Z_n^k}{n}\right| \geq \frac{3}{K+1}+3K\varepsilon\right) \leq 4Ke^{-\varepsilon^2 n/8}.$$

If we firstly change ε to $(\varepsilon/6)^2$ and then select $K := \lfloor 6/\varepsilon \rfloor$, we derive

$$\inf_{K\in\mathbf{N}}\left\{\mathbf{P}\left(\frac{3}{K+1}+3\sum_{k=1}^{K}\frac{1}{k}\left|\frac{Z_1^k+\ldots+Z_n^k}{n}\right| \geq \varepsilon\right)\right\} \leq \min\left(1,(24/\varepsilon)e^{-\varepsilon^4 n/288}\right). \tag{4.2.7}$$

By the Erdös–Turán inequality (4.1.2), the left-hand side of (4.2.7) is an upper bound on our requisite $\mathrm{vol}(B_\varepsilon^n) = \mathrm{vol}(\{D(\mathbf{t}) \geq \varepsilon\})$. The right-hand side of (4.2.7) is majorized by $100\,e^{-\varepsilon^4 n/300}$ for all $n \geq 1$ and all $\varepsilon > 0$. □

5. The cost of an integration

The basic idea of our solution [CDT21] of the unbounded denominators conjecture is that we can get useful holonomy bounds on certain $\mathbf{Q}(x)$-linear spaces of algebraic functions that come from a supposed $\mathbf{Z}[\![q]\!]$ modular function $f(\tau)$ on a noncongruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$, expanded formally in the modular function $x = \lambda/16$ via the equality of rings $\mathbf{Z}[\![q]\!] = \mathbf{Z}[\![x]\!]$. In that setting, the key point was in getting an asymptotically tight holonomy bound which only runs into a contradiction upon successively including more and more functions with the transformation $f(\tau) \rightsquigarrow f(p\tau)$ for a range of primes p, and finding that the increase in the dimension of $\mathbf{Z}[\![q]\!]$ modular functions is more than the increase in the holonomy bound, unless $f(\tau)$ was congruence to begin with.

In our present paper, we have a somewhat analogous scheme where the role of the transformation $f(\tau) \rightsquigarrow f(p\tau)$ is taken up by an integration $f(x) \rightsquigarrow \int (f(x) - f(0))\frac{dx}{x}$. Here it is more of a gamble whether or not the increase in the dimension (which we compute in § 12 and § 14.3 for our main application to Theorems A and C) turns out enough of a compensation for the increase in the bound (which comes entirely through the added denominators, and is handled in the present section by a prime number theorem estimate). We find it astonishing that the integrations gamble succeeds as a crucial ingredient for both of our main applications in the present paper: Theorem A on the $\mathbf{Q}$-linear independence of $1, \zeta(2)$, and $L(2, \chi_{-3})$, and Theorem C on the irrationality of certain products of two logarithms.

This section establishes some preparatory Lemmas which will be used to compute the added denominator cost for including such integrations into the setup of Theorem 2.5.1. The upshot will be the integration cost function of Definition 6.0.1 of the next section, and our main result Theorem 6.0.2 where this function is used to define an added denominators term $\tau^\sharp$ to the $\tau(\mathbf{b})$ of Theorem 2.5.1.

The following lemma is a direct consequence of the prime number theorem.

Lemma 5.0.1.

(1) If $k \sim \gamma n$ for a fixed $\gamma \in (0, 1]$, the lowest common multiple $L_{n,k}$ of the consecutive integers $n-k, n-k+1, \ldots, n$ is asymptotic under $n \to \infty$ to

$$\exp\left(\left(\sum_{h=1}^{\lfloor 1/\gamma \rfloor - 1} \frac{1}{h}\right) k + \frac{n}{\lfloor 1/\gamma \rfloor} + o(n)\right).$$

This bound is uniform for all $\gamma \geq \gamma_0$, where $\gamma_0 > 0$ is a fixed constant, in the following sense, for any $\epsilon > 0$, there exists $N = N(\gamma_0, \epsilon)$ such that for all $n \geq N$ and $\gamma \geq \gamma_0$, the error term is at most ϵn.

(2) If $k = o(n)$, then the lowest common multiple $L_{n,k}$ of the consecutive integers $n-k, n-k+1, \ldots, n$ is $\exp(o(n))$. Moreover, as $n \to \infty$, we have for all $0 \leq k \leq n$,

$$L_{n,k} \leq \exp\left(\left(\sum_{h=1}^{\lfloor 1/\gamma \rfloor - 1} \frac{1}{h}\right) k + \frac{n}{\lfloor 1/\gamma \rfloor} + o(n)\right),$$

where $\gamma = k/n$ and and if $k = 0, \gamma = 0$, the above formula is to be interpreted as $\exp(o(n))$. The error term $o(n)$ in this upper bound is uniform.

Basic Remark 5.0.2. If $k \geq n/2$, then $[1, 2, \ldots, n] = [(n-k), \ldots, n]$, because if $p \leq n$ then some multiple of p lies in $[n/2, n]$. Hence $L_{n,k}$ does not depend on k

within this range. With $\gamma = k/n$ and $n \to \infty$, the exponent in this inequality in terms of γ as a multiple of n is given in Figure 5.0.3. △

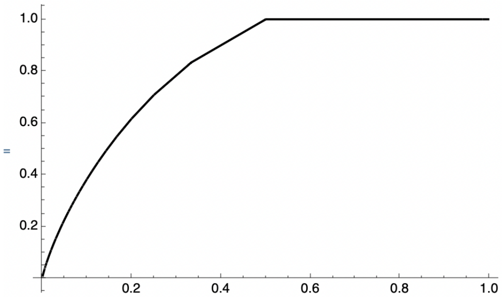

FIGURE 5.0.3. The bound for $\log(L_{n,k})/n$ as a function of $\gamma = k/n$ with $n \to \infty$.

Proof of Lemma 5.0.1. We begin with part (1). By the prime number theorem, the main term in $[1, \ldots, n]$ (after taking log) is given by $\sum_{p \le n} \log p$ (i.e., we may just count primes without counting multiplicities). The error term here is independent of γ. Thus the exponential asymptotic rate of $[n-k, \ldots, n]$ is given by counting how many of the primes $p \le n$ divide at least one among $n-k, \ldots, n$. The only primes $p \le n$ not occurring in this count are those that admit an $a \in \mathbf{N}_{>0}$ such that $ap < n-k$ and $(a+1)p > n$. Given such an a, the primes p in question are exactly the primes from the interval $(n/(a+1), (n-k)/a)$. This is a non-empty interval if and only if $a+1 < 1/\gamma$; in which case its length equals $n((1-\gamma)/a - 1/(a+1))$. Hence, $\log\,[n-k, \ldots, n]$ amounts to

$$n\left(1 - \sum_{a=1}^{\lfloor 1/\gamma \rfloor - 1} ((1-\gamma)/a - 1/(a+1))\right) + o(n)$$
$$= n\left(\gamma\left(\sum_{a=1}^{\lfloor 1/\gamma \rfloor - 1} 1/a\right) + 1/(\lfloor 1/\gamma \rfloor)\right) + o(n).$$

Note that from our assumption $\gamma \ge \gamma_0$, the above sum is a finite sum with a uniform upper bound on its number of terms. Also, the error term from the prime number theorem is uniformly controlled as it is only being applied to intervals whose lengths and endpoints are controlled uniformly in n.

We now consider case (2) of Lemma 5.0.1. The precise formulation of the first assertion is that for any $\epsilon > 0$, there exist $N = N(\epsilon)$ and $\delta = \delta(\epsilon)$ such that for all $n > N$ and all $k < \delta n$, we have $L_{n,k} < \epsilon n$. This is a consequence of (1). More precisely, for $\delta < 1/2$ by definition, for all $k < \delta n$, we have

$$L_{n,k} \le L_{n,\delta n} = \exp\left(\left(\sum_{h=1}^{\lfloor 1/\delta \rfloor - 1} \frac{1}{h}\right)\delta n + \frac{n}{\lfloor 1/\delta \rfloor} + o_\delta(n)\right).$$
$$\le \exp\left(\left(\delta(1 + \log(1/\delta)) + (1/\delta - 1)^{-1}\right) n + o_\delta(n)\right)$$

Note that $\lim_{\delta\to 0}\delta(1+\log(1/\delta))+(1/\delta-1)^{-1}=0$; we pick a $\delta=\delta(\epsilon)$ with $\delta(1+\log(1/\delta))+(1/\delta-1)^{-1}<\epsilon/2$. For this δ, by (1), there exists $N=N(\delta,\epsilon)=N(\epsilon)$ such that for $n>N$, the error term $o_\delta(n)<(\epsilon/2)n$. Then for all $n>N$ and $k<\delta n$, we have $L_{n,k}<\epsilon n$.

The precise formulation of the second assertion is that for any $\epsilon>0$, there exists $N=N(\epsilon)$ such that for all $n>N$ and all $0\le k\le n$, we have

$$\log L_{n,k}\le\left(\sum_{h=1}^{\lfloor 1/\gamma\rfloor-1}\frac{1}{h}\right)k+\frac{n}{\lfloor 1/\gamma\rfloor}+\epsilon n.$$

Note that from the proof of the first assertion above, there exists $\delta=\delta(\epsilon)$ such that the above inequality holds for all $n>N_1(\epsilon)$ and $k<\delta n$. Moreover, from part (1) with $\gamma_0=\delta$, the above inequality holds for all $n>N_2(\delta,\epsilon)=N_2(\epsilon)$. Thus the desired bound holds for all $n>\max\{N_1(\epsilon),N_2(\epsilon)\}$. □

The following is a variant of the above lemma.

Lemma 5.0.4.

(1) Fix $\gamma_0\le\gamma_0'\in(0,1)$. For $k,l\le n$ such that $\gamma_0\le k/n$ and $\gamma_0\le l/n\le\gamma_0'$, the logarithm of the product $L_{n,k}^{\ge l}$ of the primes $p>l$ that have some multiple in the interval $[n-k,n]$ is asymptotic — as $n\to\infty$ uniformly with respect to k,l — to

$$\left(k\sum_{h=1}^{\lfloor (n-k)/\max(k,l)\rfloor}1/h\right)+\left(\frac{n}{\lfloor (n+(l-k)^+)/\max(k,l)\rfloor}-l\right)^+ +o(n),$$

where $\alpha^+:=\max(0,\alpha)$ and the convention being that $\sum_{h=a}^{b}$ is over all integers in the range $a\le h\le b$, and the empty sum is zero.

(2) Moreover, as $n\to\infty$, for all $0\le k,l\le n$, we have

$$\log L_{n,k}^{\ge l}\le\left(k\sum_{h=1}^{\lfloor (n-k)/\max(k,l)\rfloor}1/h\right)+\left(\frac{n}{\lfloor (n+(l-k)^+)/\max(k,l)\rfloor}-l\right)^+ +o(n),$$

where the error term is uniform with respect to all k,l.

(If $k=l=0$, the right-hand side of the above bound is to be interpreted as $o(n)$.)

Proof. We begin with part (1). As in the proof of Lemma 5.0.1, the primes $p\le n$ that do not have any multiples among $n-k,\dots,n$ are the ones that lie in $\cup_{a=1}^{\lfloor n/k\rfloor-1}(n/(a+1),(n-k)/a)$. The new assumption here that $p>l$ implies that $a<(n-k)/l$, and hence that

$$\begin{aligned}&\left(\cup_{a=1}^{\lfloor (n-k)/k\rfloor}(n/(a+1),(n-k)/a)\right)\cap(l,n]\\&=\left(\cup_{a=1}^{h_0-1}(n/(a+1),(n-k)/a)\right)\cup(\max(n/(h_0+1),l),(n-k)/h_0),\end{aligned}$$

where $h_0=\lfloor (n-k)/\max(k,l)\rfloor$.

By the prime number theorem, our asymptotic is given by

$$n-l-\left(\sum_{a=1}^{h_0}((n-k)/a-n/(a+1))-(\max(n/(h_0+1),l)-n/(h_0+1))\right)+o(n)$$
$$=k\left(\sum_{a=1}^{h_0}1/a\right)+\max(n/(h_0+1),l)-l+o(n)=k\left(\sum_{a=1}^{h_0}1/a\right)+(n/(h_0+1)-l)^+ +o(n)$$

and $n/(h_0+1)=\dfrac{n}{\lfloor (n-k)/\max(k,l)\rfloor+1}=\dfrac{n}{\lfloor (n+(l-k)^+)/\max(k,l)\rfloor}$.

The precise formulation of the second assertion (2) is that for any $\epsilon>0$, there exists $N=N(\epsilon)$ such that for all $n>N$ and all $0\le k,l\le n$, we have

$$\log L_{n,k}^{\ge l}\le\left(k\sum_{h=1}^{\lfloor (n-k)/\max(k,l)\rfloor}1/h\right)+\left(\frac{n}{\lfloor (n+(l-k)^+)/\max(k,l)\rfloor}-l\right)^+ +\epsilon n.$$

By Lemma 5.0.1 (2), there exists $\delta=\delta(\epsilon)$ such that for all $n>N_1(\epsilon)$ and all $k\le\delta n$, we have

$$\log L_{n,k}^{\ge l}\le\log L_{n,k}<\epsilon n.$$

Therefore, we now assume $k>\delta n$. For $l<\min\{\delta,\epsilon/2\}n<k$, by Lemma 5.0.1(1), we have that for $n>N_2(\delta,\epsilon/2)=N_2(\epsilon)$,

$$\begin{aligned}\log L_{n,k}^{\ge l}\le\log L_{n,k}&\le\left(k\sum_{h=1}^{\lfloor (n-k)/k\rfloor}1/h\right)+\frac{n}{\lfloor n/k\rfloor}+(\epsilon/2)n\\&\le\left(k\sum_{h=1}^{\lfloor (n-k)/k\rfloor}1/h\right)+\left(\frac{n}{\lfloor n/k\rfloor}-l\right)^+ +\epsilon n,\end{aligned}$$

which is the desired bound as $\max(k,l)=k$ in this case. For $l>(1-\delta/2)n$, by definition, we have $\log L_{n,k}^{\ge l}\le\log L_{n,k}^{\ge(1-\delta/2)n}$. Applying (1) to $\gamma_0=\delta,\gamma_0'=1-\delta/2$, we have that there exists $N_3=N_3(\delta)=N_3(\epsilon)$ such that for all $n>N_3$, we have

$$\log L_{n,k}^{\ge(1-\delta/2)n}\le\left(k\sum_{h=1}^{\lfloor (n-k)/\max(k,(1-\delta/2)n)\rfloor}1/h\right)$$
$$+\left(\frac{n}{\lfloor (n+((1-\delta/2)n-k)^+)/\max(k,(1-\delta/2)n)\rfloor}-(1-\delta/2)n\right)^+ +(\epsilon/2)n.$$

Note that the first term is 0 since $n-k<(1-\delta/2)n$ and the second term $\le n-((1-\delta/2)n\le(\delta/2)n$. For the above proof, we may shrink δ to make it $<\epsilon$ and then the above discussion shows that for all $l\ge(1-\delta/2)n$, we have the desired bound

$$\log L_{n,k}^{\ge l}\le\log L_{n,k}\le\epsilon n.$$

Now we only remain to consider $k\ge\delta n$ and $\min\{\delta,\epsilon/2\}n\le l<(1-\delta/2)n$ and this case follows from (1). ☐

6. The fine holonomy bound

In this section, we arrive at our first main holonomy bound (Theorem 6.0.2), which we prove by revisiting the method in [CDT21, § 2.5] and enhancing it by the (standard) results of § 3 and § 4. This elementary treatment of our bound suffices for the proof of Theorems A and C and for all our other applications in this paper. Later, in § 7 and § 8, we will prove other holonomy bounds, some of which involve a Bost–Charles double integral that is theoretically smaller than the rearrangement integral in (6.0.15); however, we will find in Remark 8.1.17 and § 8.3 the difference to be negligibly small in practice. For our default treatment we have opted to highlight the increasing rearrangement feature which occurs, under a probabilistic interpretation, simultaneously in the top and bottom quantities in the holonomy quotient (6.0.10).

In order to state our holonomy bound, we first define (following Lemma 5.0.4) a function to measure up the additional contributions to the coefficient denominators in our multivariable evaluation module under including the *integrals* of our original set of functions, as discussed in § 5. (In the statement of Theorem 6.0.2, these will be the functions f_i of the form (6.0.9) with $e_i > 0$.)

Definition 6.0.1. For $0 \leq \max\{u, 1\} \leq v$ and $w \leq v$, set

$$I_u^v(w) := \int_{\min\{u,1\}}^1 \max\{t - w, 0\}\, dt + \int_{\max\{u,1\}}^v \left\{ \sum_{h=1}^{\lfloor (t-1)/\max(1,w) \rfloor} 1/h \right\} dt + \int_{\max\{u,1\}}^v \max\left\{ \frac{t}{\lfloor (t + \max(0, w-1))/\max(1,w) \rfloor} - w, 0 \right\} dt.$$

We now have:

Theorem 6.0.2. *Consider two positive integers $m, r \in \mathbf{N}_{>0}$, a nonnegative integer vector $\mathbf{e} := (e_1, \ldots, e_m) \in \mathbf{N}^m$, and an $m \times r$ rectangular array of nonnegative real numbers*

$$\mathbf{b} := \big(b_{i,j}\big)_{1 \leq i \leq m,\, 1 \leq j \leq r},$$

all of whose columns have the form

$$0 = b_{1,j} = \cdots = b_{u_j,j} < b_{u_j+1,j} = \cdots = b_{m,j} =: b_j, \qquad \forall j = 1, \ldots, r, \tag{6.0.3}$$

for some $u_j \in \{0, 1, \ldots, m\}$ depending on the column. Let

$$\sigma_i := b_{i,1} + \ldots + b_{i,r}, \qquad i = 1, \ldots, m$$

be the i-th row sum, and define

$$\tau^\flat(\mathbf{b}) := \frac{1}{m^2} \sum_{i=1}^m (2i-1)\sigma_i = \sigma_m - \frac{1}{m^2} \sum_{j=1}^r u_j^2 b_j \in [0, \sigma_m]. \tag{6.0.4}$$

and, with $I_\xi^m(\xi)$ as in Definition 6.0.1,

$$\tau^\sharp(\mathbf{e}) := (2/m^2) \min_{\xi \in [0,m]} \left\{ \xi \sum_{i=1}^m e_i + \left(\max_{1 \leq i \leq m} e_i \right) I_\xi^m(\xi) \right\}. \tag{6.0.5}$$

Define, finally,

$$\tau(\mathbf{b}; \mathbf{e}) := \tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e}). \tag{6.0.6}$$

Consider a sequence of holomorphic mappings $\varphi_0,\dots,\varphi_l : (\overline{\mathbf{D}},0) \to (\mathbf{C},0)$ *with derivatives (*conformal sizes*) satisfying*

$$|\varphi_0'(0)| < |\varphi_1'(0)| < \cdots < |\varphi_l'(0)| \quad \text{and} \quad |\varphi_l'(0)| > e^{\max(\sigma_m,\tau(\mathbf{b};\mathbf{e}))}. \tag{6.0.7}$$

Accordingly, partition the segment $[0,m]$ *by introducing the division point parameters*

$$0 = \gamma_0 < \cdots < \gamma_l < \gamma_{l+1} = m,$$

and use these choices to define an L^1 *function by piecewise patching the functions* $\log|\varphi_k|$ *on the circle* $\mathbf{T}$ *according to the linear scaling of* $[0,1)$ *to* $[\gamma_k/m, \gamma_{k+1}/m)$*:*

$$\begin{aligned} &g_{\boldsymbol{\varphi},\boldsymbol{\gamma}} : [0,1) \to \mathbf{R}\cup\{-\infty\}, \\ &g_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t) := \log\left|\varphi_k\left(e^{2\pi i \frac{mt-\gamma_k}{\gamma_{k+1}-\gamma_k}}\right)\right| \quad \text{on} \quad t\in[\gamma_k/m, \gamma_{k+1}/m)\,. \end{aligned} \tag{6.0.8}$$

If there exists an m*-tuple* $f_1,\dots,f_m \in \mathbf{Q}[\![x]\!]$ *of* $\mathbf{Q}(x)$*-linearly independent formal functions with denominator types of the form*

$$f_i(x) = a_{i,0} + \sum_{n=1}^{\infty} a_{i,n} \frac{x^n}{n^{e_i}[1,\dots,b_{i,1}\cdot n]\cdots[1,\dots,b_{i,r}\cdot n]}, \qquad a_{i,n}\in\mathbf{Z}, \tag{6.0.9}$$

such that $f_i(\varphi_k(z)) \in \mathbf{C}[\![z]\!]$ *is the germ of a meromorphic function on* $|z|<1$ *for all pairs* $i=1,\dots,m$ *and* $k=0,\dots,l$*, then*

$$\begin{aligned} m &\le \frac{\displaystyle\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt + \frac{1}{m}\sum_{k=1}^{l}\gamma_k^2\log\frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|}}{\log|\varphi_l'(0)| - \tau(\mathbf{b};\mathbf{e})} \\ &= \frac{\displaystyle\int_0^1\int_0^1 \max\left(g_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(s), g_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\right)\,ds\,dt + \frac{1}{m}\sum_{k=1}^{l}\gamma_k^2\log\frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|}}{\log|\varphi_l'(0)| - \tau(\mathbf{b};\mathbf{e})}. \end{aligned} \tag{6.0.10}$$

If moreover all functions f_i *are* a priori *assumed to be holonomic, the assumption* $|\varphi_l'(0)| > e^{\max(\sigma_m,\tau(\mathbf{b};\mathbf{e}))}$ *on* φ_l *in equation (6.0.7) can be relaxed to* $|\varphi_l'(0)| > e^{\tau(\mathbf{b};\mathbf{e})}$.

Here, in g^*, we use the notation from (2.4.1) of the *increasing rearrangement function* of g. This is why we will often refer to quantities like $\int_0^1 2t\cdot g^*(t)\,dt = \iint_{[0,1]^2}\max\left(|g(s)|,|g(t)|\right)\,ds\,dt$ as to *rearrangement integrals.*

Remark 6.0.11 (Musical Notation)**.** In the notation of Theorem 2.5.1, we have $\tau(\mathbf{b}) = \tau^\flat(\mathbf{b}) = \tau(\mathbf{b};\mathbf{0})$. Our reason for the musical notation is to think of $\tau = \tau^\flat(\mathbf{b})$ as the main reduction (flattening) of cruder values such as the value $\tau = \sigma_m$ from [CDT24] when we remove the powers $n^{\mathbf{e}}$ from (7.0.1), and of $\tau^\sharp(\mathbf{e})$ as the extra term from adding those integrations to the original list of functions. △

Remark 6.0.12. In Theorem 6.0.2 (and all the other similar theorems that we prove), we may formally relax the denominator type (6.0.9) to allow for the looser form:

$$n^{e_i}[1,\dots,b_{i,1}\cdot n + c_{i,1}]\cdots[1,\dots,b_{i,r}\cdot n + c_{i,r}], \tag{6.0.13}$$

for any fixed set of integers $c_{i,j}$. This follows upon applying the original statement of our theorem where all the nonzero $b_{i,j}$ are changed to $b_{i,j}+\varepsilon$, for some sufficiently small positive number $\varepsilon>0$. This subsumes the denominator type (6.0.13) for all but finitely many n, and any finite initial string of coefficients can be made to have any given denominator type by scaling. Then one takes the limit $\varepsilon \to 0$, after noting that the bounds always depend continuously on the $b_{i,j}$. △

As the special case $l = 0$ of a single analytic map φ, we record the extension of the bound (2.5.5):

Corollary 6.0.14. *Assume the same conditions and notation as in Theorem 6.0.2, but consider more simply a single holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ satisfying $|\varphi'(0)| > e^{\tau(\mathbf{b};\mathbf{e})}$ and such that the pullbacks $\varphi^* f_i$ are meromorphic functions on the open unit disc, that is, $\varphi^* f_i \in \mathcal{M}(\mathbf{D})$. If either $|\varphi'(0)| > e^{\sigma_m}$, or if all functions f_i are holonomic, then*

$$m \leq \frac{\iint_{\mathbf{T}^2} \log\left(\max(|\varphi(z)|, \varphi(w)|\right) \, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} = \frac{\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t})|)^* \, dt}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \tag{6.0.15}$$

Remark 6.0.16. The *a priori* holonomicity cannot be dropped if we only assume $|\varphi'(0)| > e^{\tau(\mathbf{b};\mathbf{e})}$. More precisely, for any given datum $(\mathbf{b};\varphi)$ in the statement of Corollary 6.0.14 (the case of a single map φ in Theorem 6.0.2), *except* now with assuming the opposite inequality $|\varphi'(0)| \leq e^{\sigma_m}$, a simple inductive construction demonstrates the following. If there is at least one m-tuple of $\mathbf{Q}(x)$-linearly independent formal functions $\{f_i\}$ obeying the arithmetic and analytic conditions of the datum $(\mathbf{b};\varphi)$, then there are continuum-many such m-tuples; this in particular implies non-holonomic such functions.

To see the claim, upon keeping fixed $f_1, \ldots, f_{m-1}$, it suffices to show that $f_m \in \mathbf{Q}[\![x]\!]$ has continuum-many valid coefficient options $a_{m,\bullet} \in \mathbf{Z}$ in the form (6.0.9) with $e_m = 0$, under which the pulled back power series $\sum c_n z^n := f_m(\varphi(z)) \in \mathbf{C}[\![z]\!]$ has sub-exponentially small coefficients $|c_n| = \exp(o(n))$. (Compare with [BC22, § 6.4.2], [Pól1923, § 6], or [Rob68, § 5].) We show that each successive coefficient $a_{m,n} \in \mathbf{Z}$ has at least two valid options after all the preceding coefficients $a_{m,0}$, …, $a_{m,n-1}$ have already been selected. This follows upon recursively expressing $c_n = \mu_n a_{m,n} - P_n(a_{m,0}, \ldots, a_{m,n-1})$ with coefficient

$$\mu_n := \varphi'(0)^n \Big/ \prod_{i=1}^{h} [1, \ldots, b_{m,i} \cdot n]$$

of sub-exponential growth by the prime number theorem, and $P_n \in \mathbf{C}[x_0, \ldots, x_{n-1}]$ polynomials that depend on the map φ. This gives the two distinct valid options $a_{m,n} \in \{\lfloor P_n(\boldsymbol{a}_{m,<n})/\mu_n \rfloor, \lfloor P_n(\boldsymbol{a}_{m,<n})/\mu_n \rfloor + 1\}$, and altogether a construction of a set of $f_m \in \mathbf{Q}[\![x]\!]$ with cardinality $2^{\#\mathbf{N}} = \#\mathbf{R}$. △

An essential technical feature in this section, and ultimately in the proofs of both Theorems A and C, is the term $\tau^\sharp$ accommodating added integrals to the principal denominators shape of Theorem 2.5.1. We describe this feature on a few examples.

Basic Remark 6.0.17. To revisit the simplest example from Basic Remark 2.6.3, let us compute the quantity $\tau(\mathbf{b};\mathbf{e})$ with

$$\mathbf{b} = (0,0)^{\mathrm{t}}, \qquad \mathbf{e} = (0,1).$$

Clearly $\tau^\flat(\mathbf{b}) = 0$, and we easily compute that

$$\tau^\sharp(\mathbf{e}) = \frac{1}{2} \min_{\xi \in [0,2]} \left\{ \frac{3 + (\xi - 1)^2}{2} \right\} = \frac{3}{4},$$

attained at the midpoint $\xi = 1$. Hence we find the same value $\tau(\mathbf{b},\mathbf{e}) = \tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e}) = 3/4$ as when we use the cruder scheme

$$\mathbf{b} = (0,1)^{\mathrm{t}}, \qquad \mathbf{e} = (0,0), \qquad \tau^\flat(\mathbf{b}) = 3/4, \quad \tau^\sharp(\mathbf{e}) = 0,$$

and no improvement is made over § 2.5 in this example.

In a moment, we will revisit and refine the Diophantine approximation framework of [CDT21, § 2.1]. In that framework on our running example, we have auxiliary functions (replicated to many variables) of the form $P(x) + Q(x)\log(1-x)$, where $P, Q \in \mathbf{Z}[x]$ are polynomials of degrees less than D. By the discussion in § 3.3.7, the lowest order monomial βx^n of any such function is necessarily in degree $n \leq 2D-1$, where the equality is attained uniquely by the Hermite–Padé approximants which are essentially given by Legendre polynomials. Our proof scheme combines an analytic upper bound on the coefficient $\beta \in \mathbf{Q}^\times$ with the arithmetic lower bound $|\beta| \geq 1/\mathrm{den}(\beta)$ by the reciprocal of the denominator of the nonzero rational number β. We can directly see why in this case the finer denominators of $\log(1-x) = -\sum_{k=1}^\infty x^k/k$ do not give any extra help in the arithmetic lower bound $|\beta| \geq 1/\mathrm{den}(\beta)$. The denominator of the x^n coefficient $\beta \in \mathbf{Q}$ is estimated by the lowest common multiple of all integers from the interval $[n-D, n] \supset [n/2, n]$. As $[n/2, \ldots, n] = [1, \ldots, n]$ (for every integer $k \in [1, n]$ has a unique 2-power multiple fitting into $(n/2, n]$), this in the situation is equal to the lowest common multiple $[1, \ldots, n]$ of the full initial string of integers: the estimate that we get from using $[1, \ldots, k]$ instead of k as the coefficients denominators in the function $f(x) = \log(1-x)$. We also see that the finer denominators are expected to make a difference once we have at least $m \geq 3$ functions, as already $[2n/3, \ldots, n]$ is substantially smaller than $[1, \ldots, n]$. (See also Basic Remark 5.0.2.) △

Remark 6.0.18. Using a refined pair $\mathbf{b} \in M_{m\times r}(\mathbf{N})$ with an $\mathbf{e} \in \mathbf{N}^m$ instead of a crude concatenation with $\mathbf{e} \rightsquigarrow \mathbf{0}$ may not always give an improvement in the estimate of Theorem 6.0.2. Consider the case $m = 3$ with the situation with the proof of Theorem 2.7.2, but with the finer types

$$\mathbf{b} = (0,0,1)^{\mathrm{t}}, \qquad \mathbf{e} = (0,1,0),$$

giving a template of three functions with denominator types

$$x^n, \qquad \frac{x^n}{n}, \qquad \frac{x^n}{[1,\ldots,n]}. \tag{6.0.19}$$

This choice for the array $(\mathbf{b};\mathbf{e})$ has $\tau^\flat(\mathbf{b}) = \tau^\sharp(\mathbf{e}) = 5/9$, with the latter reaching the minimum over the whole interval $\xi \in [3/2, 2]$. But the type (6.0.19) is also covered by the cruder choice

$$\mathbf{b}_0 := (0,1,1)^{\mathrm{t}}, \qquad \mathbf{e}_0 = (0,0,0),$$

that we made for the proof of Theorem 2.7.2, and this basic choice gives in this case the better value $\tau(\mathbf{b}_0;\mathbf{e}_0) = \tau(\mathbf{b}_0) = 8/9$ than $\tau(\mathbf{b};\mathbf{e}) = 10/9$. The reason for this is in how the denominators in the leading order coefficients of the auxiliary functions end up getting estimated in the proof of Theorem 6.0.2; the rate $\tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e})$ serves as an upper bound, and that upper bound estimation turns out to be strict and lossful in the example at hand. We do not know whether or not the upper bound is an equality in the $m = 14$ case that we ultimately devise for the proof of Theorem A, but we expect it to be a fairly sharp denominator estimate, and possibly an equality. △

Example 6.0.20. We give one final example, which we will use in § 6.8 to complete the proof of Theorem 2.8.4. For the case

$$\mathbf{b} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \\ 0 & 2 \\ 1 & 2 \end{pmatrix}, \qquad \mathbf{e} = (0,0,1,0)$$

relevant to a set of four functions with denominator types

$$x^n, \qquad \frac{x^n}{[1,\ldots,2n]}, \qquad \frac{x^n}{n[1,\ldots,2n]}, \qquad \frac{x^n}{[1,\ldots,n][1,\ldots,2n]}, \tag{6.0.21}$$

we have

$$\tau^{\flat}(\mathbf{b}) = \frac{37}{16}, \quad \tau^{\sharp}(\mathbf{e}) = \frac{7}{16}, \qquad \tau(\mathbf{b};\mathbf{e}) = \frac{21}{8} = 2.625,$$

with the $7/16$ value being attained on the identical interval minimizer $\xi \in [2,3]$.

But even the intermediate crude choice

$$\mathbf{b}_0 = \begin{pmatrix} 0 & 0 \\ 0 & 2 \\ 1 & 2 \\ 1 & 2 \end{pmatrix}, \qquad \mathbf{e}_0 = (0,0,0,0),$$

that minimally covers the types (6.0.21) within the framework of Theorem 2.5.1 highlighted for the introduction, already gives

$$\tau(\mathbf{b}_0;\mathbf{e}_0) = \tau(\mathbf{b}_0) = \frac{1}{16}\left(1\cdot 0 + 3\cdot 2 + 5\cdot 3 + 7\cdot 3\right) = \frac{21}{8} = 2.625.$$

In this case the value is the same, similarly to the situation in Basic Remark 6.0.17. △

In contrast to the examples above, exploiting a refined pair $(\mathbf{b};\mathbf{e})$ does often lead to strictly better results than are possible by capping up to some cruder $\mathbf{e} = \mathbf{0}$ scheme. This is in particular true for the proof Theorem A laid out in § 13. There we have a local system of rank $m = 14$ with added integrals, meaning $e_i = 1$ for six of the indices, and $e_i = 0$ for the remaining eight indices. Such a vector has $\tau^{\#}(\mathbf{e}) = 27/80$ after a simple computation (13.0.5). Overall the fine $\tau(\mathbf{b};\mathbf{e})$ used computes to $191/49 + 27/80 = 16603/3920 = 4.235459\ldots$. But with $\mathbf{e} = \mathbf{0}$ types in this example we do not get a better estimate than the rather poor $865/196 = 4.413265\ldots$ of Remark 13.0.7.

6.1. **Horizontal integration.** Our proof scheme follows precisely the $d \to \infty$ asymptotic method that we originally devised for our first solution [CDT21, route § 2.5] of the unbounded denominators conjecture. This was the idea that we dubbed a *cross-variables integration*, where the given single-variable functions $f_i(x)$ were replicated in $d \to \infty$ splitting variables to form, using the Dirichlet box principle, a nonzero auxiliary function

$$F(x_1,\ldots,x_d) := \sum_{\substack{\mathbf{i}\in\{1,\ldots,m\}^d \\ \mathbf{k}\in[0,D]^d\cap\mathbf{Z}^d}} c_{\mathbf{i},\mathbf{k}}\, x_1^{k_1}\cdots x_d^{k_d}\, f_{i_1}(x_1)\cdots f_{i_d}(x_d) \in \mathbf{Q}[\![x_1,\ldots,x_d]\!] \smallsetminus \{0\}, \tag{6.1.1}$$

with integer coefficients $c_{\mathbf{i},\mathbf{k}} \in \mathbf{Z}$ of sub-exponential asymptotic size

$$|c_{\mathbf{i},\mathbf{k}}| = \exp\left(o_{d\to\infty}(dD)\right),$$

but yet with $F(x_1,\dots,x_d)$ vanishing to almost the highest conceivable order α at $\mathbf{x}=\mathbf{0}$. Curiously enough, in this scheme the Nevanlinna characteristic growth term $\int_{\mathbf{T}} \log^+|\varphi|\,\mu_{\mathrm{Haar}}$ arose not as a circle integral *per se* (although the latter is also possible, by either [CDT21, § 2.3 or § 2.4], or by our discussion based [BC22] and [CDT24] in § 7 below); but rather — by the standard Large Deviations bound mandating[20] that the low discrepancy set $D(\mathbf{z})<\epsilon$ on the high-dimensional torus $\mathbf{T}^d$ has measure at least $1-Ce^{-c\epsilon^4 d}$ once $d \geq d_0(\epsilon)$ — from the preponderant growth rate of the pulled-back monomials $\varphi(z_1)^{k_1}\cdots\varphi(z_d)^{k_d}$. This is why we would describe such an approach as doing an integration in a cross-variables way, or "horizontally" if one pictures the dimension *versus* degree *versus* $\mathbf{T}$ (the complex analysis in any one fixed variable) aspects in the Diophantine approximation construction.

One of the main findings of the present paper is a certain combination of § 3 and § 4 which allows to actually improve the meaning here of the "highest conceivable vanishing order" α in (6.1.1). In the more rudimentary treatment in [CDT21, § 2], we had only $\alpha = mdD/e - o_{d\to\infty}(dD)$ in the parameter count for the number of linear equations to be solved in the unknown variables $c_{\mathbf{i},\mathbf{k}}$; this owes to the $e^d:1$ asymptotic volumes ratio for a standard high-dimensional simplex to its largest embedded subcube. We now explain how Corollary 3.1.11, on the commutativity with Cartesian products of the formation of the vanishing filtration jumps sets, and the measure concentration material § 4 work together to improve this vanishing order to $\alpha = mdD/2 - o_{d\to\infty}(dD)$. For simplicity, since we will anyway need this later on for the general form of Theorem 6.0.2, we assume the strongest form of the Cartesian power structure (based on the holonomicity of the f_i): Lemma 3.2.14, stemming from the Chudnovsky–Osgood theorem coupled to Corollary 3.1.11.

6.1.2. *The Thue–Shidlovsky idea.* This Lemma 3.2.14 ensures that the nonzero power series (6.1.1) has to posses a nonzero monomial $\beta\,\mathbf{x}^{\mathbf{n}} = \beta\,x_1^{n_1}\cdots x_d^{n_d}$ with $\mathbf{n}=(n_1,\dots,n_d)\in[0,(m+\delta)D]^d$, for any $\delta>0$ and d, once $D \gg_{\delta,d} 1$. Therefore, in the linear system to solve for the total vanishing order in the Thue–Siegel lemma, we need not be concerned with the broad simplex region $|\mathbf{n}|<\alpha$, but instead we can simply vanish the coefficients of $\mathbf{x}^{\mathbf{n}}$ for all $\mathbf{n}$ ranging over the hypercube $[0,(m-\delta)D]^d$ (clearly, this is the hypercube of the maximal conceivable size in the parameter count), as well as for all $\mathbf{n}$ outside of the low discrepancy part of the slightly bigger hypercube $[0,(m+\delta)D]^d$ (allowing us to also use the measure concentration for the component set of the vector $\mathbf{n}$). It is essential here for the application of Theorem 4.2.1 to take $\delta>0$ sufficiently small in terms of the discrepancy parameter ϵ; then the parameter count goes through. (This step is contained in Lemma 6.2.6.) The upshot is that in this construction, as $\epsilon\to 0$ (eventually: after $D\to\infty, d\to\infty, \delta\to 0$), all the lowest order monomials $\beta\cdot x_1^{n_1}\cdots x_d^{n_d}$ in $F(x_1,\dots,x_d)$ have their exponent vectors $(n_1,\dots,n_d)$ asymptotically close to some ordering of the set $\{jmD/d : 0\leq j<d\}$. In particular, the total vanishing order is indeed close to $mdD/2$.

It is this improvement over the mdD/e of [CDT21, Lemma 2.1.2] that recovers the $e \rightsquigarrow 2$ coefficient refinement under the elementary asymptotic framework of [CDT21, § 2]. At this point, the Thue–Shidlovsky argument further supplies

[20]This is just a restatement of Theorem 4.2.1 which we treated in detail in § 4. Indeed, the exponential function establishes an isomorphism of the measure spaces $([0,1]^d,\mu_{\mathrm{Lebesgue}})$ and $(\mathbf{T}^d,\mu_{\mathrm{Haar}})$, under which the box discrepancy functions correspond.

two similar, and equally crucial, improvements on both the top and bottom of the fraction (2.2.3).

We next introduce these two improvements in turn, showcasing to the case of Corollary 6.0.14 for simplicity, following which we start on the proof of the general Theorem 6.0.2.

6.1.3. *The archimedean sharpening.* Firstly, this asymptotic scheme allows by elementary methods to improve the $\int_{\mathbf{T}} \log^+ |\varphi| \, \mu_{\mathrm{Haar}}$ integral to the strictly smaller quantity $\int_0^1 t \cdot (\log |\varphi(e^{2\pi i t})|)^* \, dt$. We have remarked already in [CDT21, § 2.3.3] that some improvement in the holonomy bound can be made by exploiting that, by Theorem 4.2.1 again, the monomials exponents $\mathbf{k}$ in (6.1.1) can be constricted to the low discrepancy part of the hypercube $[0, D]^d$. In concrete heuristic terms, this means that as $d \to \infty$, the exponents vectors $(k_1, \ldots, k_d)$ in (6.1.1) can be considered as being close to some ordering of the set $\{jD/d \, : \, 0 \leq j < d\}$. In this way, upon noting that the largest value of $\left|\varphi(z_1)^{k_1} \cdots \varphi(z_d)^{k_d}\right|$ over all these orderings occurs when $(k_1, \ldots, k_d)$ is arranged in the same way as $(|\varphi(z_1)|, \ldots, |\varphi(z_d)|)$, we find the rearrangement integral

$$\int_0^1 t \, (\log |\varphi|)^*(t) \, dt = \frac{1}{2} \iint_{\mathbf{T}^2} \log \left(\max(|\varphi(z)|, |\varphi(w)|) \right) \, \mu_{\mathrm{Haar}}(z) \mu_{\mathrm{Haar}}(w) \quad (6.1.4)$$

precisely in the refined preponderant growth rate based on the uniform distribution of not only the torus points $\mathbf{z} \in \mathbf{T}^d$, but also of the monomials exponent vector $\mathbf{k}$. An illustration of the saving thus made is in the explicit example of Figure 8.1.15.

6.1.5. *The arithmetic sharpening.* To the uniform distribution of the leading order jet exponents $\mathbf{n}$ of $F(x_1, \ldots, x_d)$, we can add yet another elementary application of the Law of Large Numbers: without changing (broadly speaking) the asymptotic size of the parameter count, we can insist that the m function species occur with equal frequency $1/m$ in all the split-variable products $f_{i_1}(x_1) \cdots f_{i_d}(x_d)$ in the make up of (6.1.1). In other words, we can assume that $d \equiv 0 \mod m$ and that the summation multi-index $\mathbf{i}$ in (6.1.1) is constricted to have each index $i_0 \in \{1, \ldots, m\}$ arise d/m times as a component of $\mathbf{i}$. This permits us to integrate over the m different denominator types of our m function species, under our condition on the denominators cap array $\mathbf{b}$, and we find precisely (2.5.6) as the counterpart of (2.4.2).

6.2. **The auxiliary construction.** We start here the proof of Theorem 6.0.2. In the case that all f_i are *a priori* holonomic functions, we have the following improvement on [CDT21, Lemma 2.1.2].

For $d \in \mathbf{N}_{>0}$ and $\epsilon \in (0, 1]$, we denote by

$$P_\epsilon^d := \left\{ \mathbf{t} \in [0,1]^d \, : \, D(\mathbf{t}) < \epsilon \right\} \subset [0,1]^d \quad (6.2.1)$$

the ϵ-discrepancy part of the d-dimensional hypercube, with $D : [0,1]^d \to [0,1]$ being the normalized discrepancy function of Definition 4.1.1. The image of this set under the analytic isomorphism $\exp : [0,1)^d \to \mathbf{T}^d$, $\mathbf{t} \mapsto e^{2\pi i \mathbf{t}}$ will be denoted by $T_\epsilon^d \subset \mathbf{T}^d$. In these notations, a cruder form of Theorem 4.2.1 can be restated as the double limits

$$\lim_{\varepsilon \to 0} \lim_{d \to \infty} \mu_{\mathrm{Lebesgue}}(P_\varepsilon^d) = 1, \qquad \lim_{\varepsilon \to 0} \lim_{d \to \infty} \mu_{\mathrm{Haar}}(T_\varepsilon^d) = 1. \quad (6.2.2)$$

In the following, we fix an $m \in \mathbf{N}_{>0}$ and restrict the asymptotic parameter $d \in \mathbf{N}_{>0}$ to the integers $\equiv 0 \mod m$. For the plan outlined in § 6.1.5, we restrict the multi-index $\mathbf{i}$ in (6.1.1) to the equidistributed set

$$V_m^d := \{\mathbf{i} \,:\, \forall i \in \{1,\ldots,m\},\ \#\{h \in \{1,\ldots,d\} \,:\, i_h = i\} = d/m\} \subset \{1,\ldots,m\}^d. \tag{6.2.3}$$

This set still has the asymptotically full size $m^{d-o(d)}$:

Lemma 6.2.4. *Under our standing assumption $d \equiv 0 \mod m$, we have*

$$\#V_m^d = \binom{d}{\frac{d}{m},\ldots,\frac{d}{m}} > m^d \Big/ \binom{d+m-1}{m-1}. \tag{6.2.5}$$

Proof. The m-fold expansion

$$m^d = (1+\ldots+1)^d = \sum_{\substack{\mathbf{j}\in\mathbf{N}^m \\ |\mathbf{j}|=m}} \binom{d}{j_1,\ldots,j_m}$$

has $\binom{d+m-1}{m-1}$ terms, the maximal of which is the central multinomial coefficient with $j_1 = \cdots = j_m = d/m$. ☐

Our Thue–Siegel construction (step (ii) of the general outline from § 2.12) is the following.

Lemma 6.2.6. *Suppose we have m* holonomic *power series $f_1,\ldots,f_m \in \mathbf{Q}[\![x]\!]$. Assume that each $f_i(x)$ has denominators of the crude type*

$$f_i(x) = \sum_{n=0}^{\infty} a_{i,n} \frac{x^n}{A^{n+1}[1,\ldots,Bn]^{\sigma}}, \qquad \textit{for some } A \in \mathbf{N}_{>0},\ B,\sigma \in \mathbf{N} \tag{6.2.7}$$

and converges on a complex disc $|x| < \rho$, for some $\rho \in (0,1)$.

There exists a function

$$d_0 : \mathbf{N}^3 \times (0,1) \times (0,1) \to \mathbf{N}$$

such that the following holds.

For each $\epsilon \in (0,1)$, there is a $\delta = \delta(\epsilon) \in (0,\epsilon)$, such that for all $d \geq d_0(A,B,\sigma,\rho;\epsilon)$ with $d \equiv 0 \mod m$, there exists asymptotically for $D \to \infty$ a nonzero d-variate formal function $F(\mathbf{x})$ of the (6.1.1) form

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{i}\in V_m^d \\ \mathbf{k}/D \in P_\epsilon^d \cap \mathbf{Z}^d}} c_{\mathbf{i},\mathbf{k}}\, \mathbf{x}^{\mathbf{k}} \prod_{s=1}^{d} f_{i_s}(x_s) \in \mathbf{Q}[\![\mathbf{x}]\!] \smallsetminus \{0\}, \tag{6.2.8}$$

with $c_{\mathbf{i},\mathbf{k}} \in \mathbf{Z}$ integers, all bounded in absolute value by $|c_{\mathbf{i},\mathbf{k}}| < e^{\epsilon d D}$, and such that the power series expansion $F(\mathbf{x}) = \sum b_{\mathbf{n}} \mathbf{x}^{\mathbf{n}}$ of (6.2.8) obeys the following main requirement:

$$(\star) \qquad \begin{array}{l} \textit{All the minimal order monomials } \beta_{\mathbf{n}}\, \mathbf{x}^{\mathbf{n}} \textit{ in } (6.2.8) \\ \textit{have an exponent vector } \mathbf{n} \textit{ satisfying } \mathbf{n}/\left((m+\delta)D\right) \in P_\epsilon^d. \end{array} \tag{6.2.9}$$

Note that $(\star)$ implies in particular that $\beta_{\mathbf{n}} \neq 0$ for at least one such $\mathbf{n}$ in the set $((m+\delta)D)\, P_\epsilon^d \cap \mathbf{Z}^d$.

Before we proceed with the proof, we collect some consequences of our condition $(\star)$. In the following, we will consider $\epsilon \in (0, 1/4]$ which in the end will be let to approach zero. Throughout this § 6, we will write

$$\alpha := mdD/2. \tag{6.2.10}$$

This notation reflects a related use of α as a vanishing parameter in [CDT21, § 2], see for example [CDT21, Lemma 2.1.2(1)]. In that previous paper, we (asymptotically) took α to be (of order) $mdD/e \sim m(d!)^{1/d}D$, which here we improve to $mdD/2$.

Corollary 6.2.11. *In Lemma 6.2.6, we have*

$$\mathrm{ord}_{\mathbf{x}=\mathbf{0}} F(\mathbf{x}) \in [(1-2\epsilon)\alpha, (1+2\epsilon)^2\alpha]. \tag{6.2.12}$$

Every multi-index $\mathbf{k} = (k_1, \ldots, k_d)$ *in* (6.2.8) *has* $\max_{j=1}^d k_j \le \frac{2\alpha}{md}$ *and admits a permutation* $\psi = \psi_{\mathbf{k}}$ *of* $\{1, \ldots, d\}$ *such that* $k_{\psi(1)} \le \cdots \le k_{\psi(d)}$ *and*

$$\frac{2\alpha j}{md^2} - 2\epsilon\alpha/(md) \le k_{\psi(j)} \le (1+\epsilon)\,\frac{2\alpha j}{md^2} + 4\epsilon\alpha/(md), \qquad \forall j \in \{1, \ldots, d\}. \tag{6.2.13}$$

Further, every exponent $\mathbf{n} = (n_1, \ldots, n_d)$ *of minimal total order*

$$n := |\mathbf{n}| = \mathrm{ord}_{\mathbf{x}=\mathbf{0}} F(\mathbf{x})$$

in the Taylor series of $F(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!]$ *has a permutation* π *of* $\{1, \ldots, d\}$ *such that* $n_{\pi(1)} \le \cdots \le n_{\pi(d)}$ *and*

$$2\alpha j/d^2 - 2\epsilon\alpha/d \le n_{\pi(j)} \le (1+\epsilon)2\alpha j/d^2 + 4\epsilon\alpha/d, \qquad \forall j \in \{1, \ldots, d\}, \tag{6.2.14}$$

and for all $u, v \in [0, 1]$ *with* $u \le v$ *it satisfies*

$$(1-2\epsilon)(v^2-u^2)\alpha \le \sum_{ud \le j < vd} n_{\pi(j)} \le (1+2\epsilon)^2(v^2-u^2)\alpha. \tag{6.2.15}$$

Proof. The partial degrees bound (6.2.13) is tautologically a rewriting of our definition (6.2.10), but it is used to organize the analysis around the leading asymptotic parameter α. The estimate (6.2.12) follows from (6.2.9) upon noting that the expected value $\mathbf{E}\,[t \in [0,1]] = \int_0^1 t\,dt = 1/2$, which by the Koksma–Hlawka inequality or an elementary bit of computation shows the implication

$$\mathbf{t} \in P_\epsilon^d \quad \Longrightarrow \quad \sum_{j=1}^d t_j \in [d/2 - \epsilon d, d/2 + \epsilon d]\,. \tag{6.2.16}$$

In detail, the upper bound in (6.2.12) is by the chain of trivial estimates $|\mathbf{n}| \le (m+\delta)D(d/2+\epsilon d) \le (1+\epsilon)mD(d/2+\epsilon d) < (1+2\epsilon)mD(d/2+\epsilon d) = (1+2\epsilon)(2\alpha/d)(d/2+\epsilon d) = (1+2\epsilon)^2\alpha$ implied by (6.2.16) for all the nonzero monomials $\beta_{\mathbf{n}}\mathbf{x}^{\mathbf{n}}$ (which form a nonempty set!) in $(\star)$. The lower bound is similar with using $|\mathbf{n}| \ge (m+\delta)D(d/2-\epsilon d) > mD(d/2-\epsilon d) = (2\alpha/d)(d/2-\epsilon d) = (1-2\epsilon)\alpha$ from (6.2.16) for all nonzero monomials $\beta_{\mathbf{n}}\mathbf{x}^{\mathbf{n}}$ in $(\star)$, and the proof of (6.2.15) is the same.

For an arbitrary $\mathbf{n}$ from the leading order $|\mathbf{n}| = \mathrm{ord}_{\mathbf{x}=\mathbf{0}} F(\mathbf{x})$ jet $(\star)$, take a permutation π with $n_{\pi(1)} \le n_{\pi(2)} \le \cdots \le n_{\pi(d)}$. For the lower bound in (6.2.14), remark that the interval $[0, n_{\pi(j)}]$ contains at least the j elements $n_{\pi(1)}, \ldots, n_{\pi(j)}$ of the component set $\{n_j\}$. As $\mathbf{n} \in ((m+\delta)D)\,P_\epsilon^d$, the definition of discrepancy mandates that the interval $[0, n_{\pi(j)}] = ((m+\delta)D) \cdot \big[0, n_{\pi(j)}/((m+\delta)D)\big)$ contains at most

$$n_{\pi(j)}d/((m+\delta)D) + \epsilon d < n_{\pi(j)}d/(mD) + \epsilon d = n_{\pi(j)}d^2/(2\alpha) + \epsilon d$$

of the components of $\mathbf{n}$. Hence

$$j \leq n_{\pi(j)} d^2/(2\alpha) + \epsilon d,$$

giving the claimed lower bound on $n_{\pi(j)}$.

For the upper bound in (6.2.14), the interval

$$\big[n_{\pi(j)}, ((m+\delta)D)\big] = ((m+\delta)D) \cdot \big[n_{\pi(j)}/\left((m+\delta)D\right), 1\big]$$

contains at least the $d-j+1$ elements $n_{\pi(j)}, \ldots, n_{\pi(d)}$ of the component set $\{n_j\}$ of our $\mathbf{n} \in ((m+\delta)D)\, P_\epsilon^d$. Thus, as before,

$$\begin{aligned} d-j+1 &\leq d \cdot \big(1 - n_{\pi(j)}/\left((m+\delta)D\right)\big) + \epsilon d \\ \leq d - n_{\pi(j)} d/\left((1+\epsilon)mD\right) + \epsilon d &= d - n_{\pi(j)}(1+\epsilon)^{-1} d^2/(2\alpha) + \epsilon d \end{aligned}$$

completing the proof of (6.2.14). Finally, the bounds (6.2.13) follow by the same proof. □

6.3. **The box principle step: proof of Lemma 6.2.6.** Our proof combines the classical Thue–Siegel lemma [BG06, Lemma 2.9.1] and Lemma 3.2.14 on the functional bad approximability combined with the product structure of the vanishing filtration jumps, following the plan we laid out in § 6.1.2.

Proof of Lemma 6.2.6. The first parameter to consider is the ϵ that we use to measure the discrepancy from $\mu_{\mathrm{Lebesgue},[0,1)}$ or μ_{Haar}; this will be the last parameter which we let approach 0 in the proof. Then, in terms of the exponential decay rate function $\kappa(\epsilon) := \epsilon^4/300 > 0$ in Theorem 4.2.1, we take any $\delta \in (0,\epsilon)$ small enough to have

$$\frac{m-\delta}{m+\delta} > e^{-\kappa(\epsilon)} = e^{-\epsilon^4/300}. \tag{6.3.1}$$

We then set up a linear system of $M \leq 2\left((m-\delta)D\right)^d$ linear equations in the $N = (m - o_{d\to\infty}(1))D)^d$ unknown coefficients $c_{\mathbf{i},\mathbf{k}}$ in the function template form (6.2.8), by requiring that in the Taylor series expansion

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{i} \in V_m^d \\ \mathbf{k}/D \in P_\epsilon^d \cap \mathbf{Z}^d}} c_{\mathbf{i},\mathbf{k}}\, \mathbf{x}^{\mathbf{k}} \prod_{s=1}^d f_{i_s}(x_s) = \sum_{\mathbf{n} \in \mathbf{N}^d} \beta_{\mathbf{n}}\, \mathbf{x}^{\mathbf{n}}, \tag{6.3.2}$$

all coefficients $\beta_{\mathbf{n}}$ vanish whenever either $\max_{j=1}^d n_j \leq (m-\delta)D$ or $\mathbf{n} \notin (m+\delta)D \cdot P_\epsilon^d$. Once the dimension $d \gg_{\epsilon,\delta} 1$ is sufficiently big, Theorem 4.2.1 and Lemma 6.2.4 show that the number N of free parameters $c_{\mathbf{i},\mathbf{k}}$ in our linear system will exceed the quantity

$$N > \big((m-\delta/2)D\big)^d > 2\big((m-\delta)D\big)^d \geq M. \tag{6.3.3}$$

In the Siegel lemma, this gives a *Dirichlet exponent*

$$\frac{M}{N-M} < \frac{1}{\frac{1}{2} \cdot \left(\frac{m-\delta/2}{m-\delta}\right)^d - 1} = o_{d\to\infty}(1). \tag{6.3.4}$$

For the height of our linear system, a simple estimate based on the prime number theorem shows that the system can be expressed into the form $\mathbf{A} \cdot \mathbf{y} = 0$, to be solved nontrivially for an integer vector $\mathbf{y} \in \mathbf{Z}^N$ of a small height, with some $M \times N$ integer matrix $\mathbf{A} \in \mathrm{M}_{M\times N}(\mathbf{Z})$ whose entries are bounded in absolute value by $C_0(A,B,\sigma,\rho)^\alpha$. Here, $C_0(A,B,\sigma,\rho)$ is a simple computable function, immaterial to us, in the parameters A, B, σ, ρ that we assume for the form (6.2.7) of the functions $f_i(x)$ (archimedeanly convergent on $|x| < \rho$). At this point (6.3.4) and [BG06, Lemma 2.9.1] prove that, once $d \gg_{\epsilon,\delta} 1$ and then $D \gg_d 1$, there exists a nonzero formal function $F(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!] \smallsetminus \{0\}$ of the form (6.3.2), in which all coefficients $c_{\mathbf{i},\mathbf{j}} \in \mathbf{Z}$ on the left-hand side are rational integers smaller than $e^{\epsilon dD}$ in absolute value, and having on the right-hand side the vanishing of all $\beta_{\mathbf{n}} = 0$ with $\mathbf{n} \notin (m+\delta)D \cdot P_\epsilon^d$, as well as for all $\mathbf{n} \notin [0,(m-\delta)D]^d$.

The desired property $(\star)$ (see Equation 6.2.9) now follows by Lemma 3.2.14, applied with $\varepsilon := \delta/2$, after noting our assumption that $f_1, \ldots, f_m$ are holonomic functions. ☐

The condition in Lemma 6.2.6 that all f_i are holonomic functions is met by the hypotheses in Theorem 6.0.2 currently under proof. Indeed, an *a priori* holonomicity is either directly an assumption, or else the stronger positivity condition (6.0.7) is imposed. By André's holonomicity criterion (Corollary 2.6.1, also outlined in § 2.12, and completely proved in the self-contained § B), upon applying to $f_i(x)$ the differential operator $\left(x\frac{d}{dx}\right)^{e_i}$ to remove the extra n^{e_i} terms from the denominators of (6.0.9) (and observing that, thanks to the chain rule, the d/dx derivation preserves the meromorphicity condition $\varphi_l^* f_i \in \mathcal{M}(\overline{\mathbf{D}})$), the condition $\log|\varphi_l'(0)| > \sigma_m \geq b_{i,1} + \ldots + b_{i,r}$ by itself forces f_i to be a holonomic function.

This places us into a position to apply Lemma 6.2.6. In the following, we will fix an "auxiliary function" $F(\mathbf{x}) \in \mathbf{Q}[\![\mathbf{x}]\!] \smallsetminus \{0\}$ supplied by that lemma, and write $\delta := \delta(\epsilon)$ for the $\delta \in (0,\epsilon)$ under the thesis of the lemma. At the end of the proof we will let, in this order, $D \to \infty$, $d \to \infty$, and $\epsilon \to 0$, remembering that the latter also in particular makes $\delta \to 0$.

6.4. **Seeding.** Consider now a nonzero minimal order monomial $\beta\,\mathbf{x}^{\mathbf{n}}$ in $F(\mathbf{x})$. Thus $\beta := \beta_{\mathbf{n}} \in \mathbf{Q}^\times$ is a nonzero rational number of a certain denominator cap inherited from (6.0.9), that we will study in § 6.6 below, and the exponent $\mathbf{n} = (n_1, \ldots, n_d) \in ((m+\delta)D)\, P_\epsilon^d$ with $\delta < \epsilon$ has $n := |\mathbf{n}| = n_1 + \ldots + n_d \in [(1-2\epsilon)\alpha, (1+2\epsilon)^2\alpha]$ by Corollary 6.2.11. *Until the end of the proof, we fix this minimal order exponent* $\mathbf{n}$*, and then upon relabeling the variables* $x_1, \ldots, x_d$*, we may and will assume that* $n_1 \leq n_2 \leq \cdots \leq n_d$.

We turn now to the piece of the argument — which we omitted from the introductory sketch § 6.1 (but we briefly described in § 2.13.10 of our general introduction), — needed to get the stronger bound (6.0.8) in place of the more basic special case (6.0.15). The idea is to partition the indexing set $\{1, \ldots, d\}$ into $l+1$ groups so as to use the map φ_k in the analytic variable z_j for the case $\gamma_k/m \leq j/d < \gamma_{k+1}/m$, for $k = 0, \ldots, l$, with the understanding that $\gamma_{l+1}/m = 1$ and equality is meant on the right-hand side condition for $k = l$. Let $\Phi : \overline{\mathbf{D}}^d \to \mathbf{C}^d$ be the diagonal map thus defined from using $\varphi_k(z_j)$ for its j^{th} coordinate function, where $k = k(j) \in \{0, \ldots, l\}$ is uniquely determined by $j \in \{\lceil \gamma_k d/m \rceil, \ldots, \lceil \gamma_{k+1} d/m \rceil - 1\}$ (and $k = l$ for $j = d$). This is a holomorphic mapping with $\Phi(\mathbf{0}) = \mathbf{0}$, and — clearing a common holomorphic denominator for the meromorphic functions $f_i(\varphi_k(z))$, $1 \leq i \leq m$, $0 \leq k \leq l$, — there is a holomorphic function $h \in \mathcal{O}(\overline{\mathbf{D}})$ with $h(0) = 1$ and

$$G(\mathbf{z}) := h(z_1) \cdots h(z_d) \cdot (\Phi^* F)(\mathbf{z}) \in \mathcal{O}\left(\overline{\mathbf{D}}^d\right) \tag{6.4.1}$$

holomorphic on some neighborhood of the closed unit polydisc. By construction, the $\mathbf{z}^{\mathbf{n}}$ coefficient of $G(\mathbf{z})$ equals[21]

$$[\mathbf{z}^{\mathbf{n}}]\,\{G(\mathbf{z})\} = \beta \cdot \exp\left(\sum_{k=0}^{l}\left(\sum_{\frac{d\gamma_k}{m}\leq j<\frac{d\gamma_{k+1}}{m}}{}^{\prime} n_j\right)\log\varphi_k'(0)\right), \tag{6.4.2}$$

where the dash in the inner summation is to remind us that for $k = l$ the term $j = d$ is supposed to also be included into the sum. By Lemma 6.2.11, the inner sum over j satisfies

$$(1-2\epsilon)(\gamma_{k+1}^2-\gamma_k^2)\frac{\alpha}{m^2} \leq \sum_{\frac{d\gamma_k}{m}\leq j<\frac{d\gamma_{k+1}}{m}} n_j \leq (1+2\epsilon)^2(\gamma_{k+1}^2-\gamma_k^2)\frac{\alpha}{m^2}. \tag{6.4.3}$$

6.5. **Equidistribution.** There are at least two ways [CDT21, § 2.4 or § 2.5] to handle the archimedean growth term in a manner compatible with our finer analysis. The Vandermondian damping factors of [CDT21, § 2.5] are based directly on the Cauchy formula, and are more in line with the cross-variables integration technique that we exploit to carry out § 6.1.3 and § 6.1.5. The Poisson–Jensen method [CDT21, § 2.4] is based on a lexicographical induction lemma [CDT21, Lemma 2.4.1] suggested to us by André; this approach is closer in spirit to our treatment in § 8. The third proof [CDT21, § 2.3] of our original holonomicity theorem for the solution of the unbounded denominators conjecture does not seem to apply to the present refinement.

Our choice hence will be to stick to the Vandermondians method for the details of the current section (nevertheless referring to [CDT21, § 2.5.1] for some of basic and well-known facts of potential theory).

6.5.1. *Vandermondians.* To set up our damping factor, we collect here some basic facts from the logarithmic potential theory in the complex plane. Given (a block of) variables $\mathbf{z} = (z_1, \ldots, z_d)$, we define

$$V(\mathbf{z}) := \prod_{i<j}(z_i - z_j) = \det\begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{d-1} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{d-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & z_d & z_d^2 & \cdots & z_d^{d-1}\end{bmatrix} \in \mathbf{Z}[z_1,\ldots,z_d] \smallsetminus \{0\}. \tag{6.5.2}$$

As in [CDT21, § 2.5.1], we note:

Lemma 6.5.3 (Fekete)**.** *The supremum of $|V(\mathbf{z})| = \prod_{1\leq i<j\leq d}|z_i - z_j|$ over the unit polycircle $\mathbf{z} \in \mathbf{T}^d$ is equal to $d^{d/2}$, with equality if and only if the points $z_1, \ldots, z_d$ are the vertices of a regular d-gon.*

Lemma 6.5.4 (Bilu)**.** *There are functions $c(\varepsilon) > 0$ and $d_0(\varepsilon) \in \mathbf{R}$ such that, for every $\varepsilon \in (0, 1]$, if $d \geq d_0(\varepsilon)$ and $\mathbf{z} = (z_1, \ldots, z_d) \in \mathbf{T}^d$ is a d-tuple with discrepancy $D(\mathbf{z}) \geq \varepsilon$, then*

$$|V(\mathbf{z})| = \prod_{1\leq i<j\leq d}|z_i - z_j| < e^{-c(\varepsilon)d^2}. \tag{6.5.5}$$

[21] Formally exponentiating the additive notation, choosing any branch for the logarithm.

Proof. See the proof in [CDT21, Lemma 2.5.8], which in turn is closely based on Bombieri and Gubler's treatment [BG06, page 103] of Bilu's equidistribution theorem for points of small canonical height on linear algebraic tori. □

At this point, we fix an $\epsilon > 0$ and a $\delta \in (0, \epsilon)$ until the end of the proof, and we assume $d \geq d_0(\epsilon, \delta)$.

6.5.6. *Holomorphic dampener.* We suppose now the $\{1, \dots, d\}$ partitioning into $l+1$ consecutive blocks from § 6.4, and we write $\mathbf{z} = \left(\mathbf{z}^{(0)}, \dots, \mathbf{z}^{(l)}\right)$ for the corresponding variable blocks. Explicitly, $\mathbf{z}^{(k)}$ enlists, in increasing labels, the variables z_j where $\gamma_k d/m \leq j < \gamma_{k+1} d/m$, and the end term $j = d$ is assumed to be included in the case $k = l$. Following [CDT21, § 2.5.14], we will dampen the integrand in the Cauchy integral formula for the coefficient (6.4.2) by using the following choice of multivariable holomorphic multiplier:

$$W(\mathbf{z}) := \prod_{k=0}^{l} V\left(\mathbf{z}^{(k)}\right)^M \in \mathbf{Z}[\mathbf{z}] \smallsetminus \{0\}, \tag{6.5.7}$$

where M is a large integer parameter to be selected in the proof.

6.5.8. *Cross-integration.* The idea for the cross-integration is simple. Consider $\mathbf{z} \in \mathbf{T}^d$ on the high-dimensional unit torus. If one of our $k \in \{0, \dots, l\}$ blocks has discrepancy $D\left(\mathbf{z}^{(k)}\right) \geq \epsilon > 0$, Lemma 6.5.4 tells us that the corresponding term $V\left(\mathbf{z}^{(k)}\right)$ in (6.5.7) is uniformly[22] exponentially small in $-d^2$. This, in combination with Lemma 6.5.3 used as a uniform upper bound on the other factors $V\left(\mathbf{z}^{(q)}\right)$ for $q \in \{0, \dots, l\} \smallsetminus \{k\}$, entails that the overall damping factor $W(\mathbf{z})$ decays at the exponential rate $-Md^2$, uniformly in $d \in \mathbf{N}_{>0}$ and $\left\{\mathbf{z}^{(q)} : q \in \{0, \dots, l\} \smallsetminus \{k\}\right\}$. This proves that

$$\sup_{\substack{\mathbf{z} \in \mathbf{T}^d \\ \exists k \colon D(\mathbf{z}^{(k)}) \geq \epsilon}} \{|W(\mathbf{z})|\} < e^{-c'(\epsilon) M d^2}, \tag{6.5.9}$$

with some function $c'(\epsilon) > 0$ depending on ϵ but not on d.

In addition, momentarily using $d = d_0 + \ldots + d_l$ to denote the partition of the variable slot cardinalities, Lemma 6.5.3 also implies the uniform $\exp(o(Md^2))$ upper bound

$$\sup_{\mathbf{z} \in \mathbf{T}^d} |W(\mathbf{z})| \leq \prod_{k=0}^{l} d_k^{M d_k / 2} \leq d^{Md/2}. \tag{6.5.10}$$

The effect of using a multiplier with (6.5.9) and (6.5.10) is roughly the following. Since the monomial exponent vectors $\mathbf{k}$ in the make up of G via (6.2.8) have asymptotically uniformly distributed components $\{k_j\} \subset [0, D]$, the rearrangement inequality brings out the function (6.0.8) and entails in the limit for the product $W(\mathbf{z})G(\mathbf{z})$ to sift out the mean growth rate

$$\exp\left(D \int_0^1 t \cdot (g_{\boldsymbol{\varphi}, \boldsymbol{\gamma}})^*(t)\, dt\right)$$

as a uniform $\mathbf{z} \in \mathbf{T}^d$ supremum. We will make this into a precise argument below.

[22] As a function of $d \in \mathbf{N}_{>0}$, but for the fixed $\epsilon > 0$.

With this plan in mind, we turn to step (iii) of the general outline from § 2.12. We study analytically the coefficient $\beta \in \mathbf{Q}^\times$ of $\mathbf{x}^\mathbf{n}$ in $F(\mathbf{x})$. To approach it, we express (6.4.2) by a Cauchy integral:

$$\begin{aligned}
&\beta \cdot \exp\left(\sum_{k=0}^{l} \left(\sideset{}{'}\sum_{\frac{d\gamma_k}{m} \le j < \frac{d\gamma_{k+1}}{m}} n_j \right) \log \varphi_k'(0) \right) \\
&= [\mathbf{z}^\mathbf{n}] \{G(\mathbf{z})\} \\
&= \left[z_1^{n_1+M} \cdots z_d^{n_d+dM} \right] \{W(\mathbf{z})G(\mathbf{z})\} \\
&= \int_{\mathbf{T}^d} \frac{W(\mathbf{z})G(\mathbf{z})}{z_1^{n_1+M} z_2^{n_2+2M} \cdots z_d^{n_d+dM}} \, \mu_{\mathrm{Haar}}(\mathbf{z}).
\end{aligned} \tag{6.5.11}$$

Consequently, estimating the latter integrand pointwise by the supremum, we derive an analytic upper bound on the nonzero rational number $\beta \in \mathbf{Q}^\times$:

$$\begin{aligned}
|\beta| &\le \exp\left(\sup_{\mathbf{T}^d} \{\log|WG|\} - \sum_{k=0}^{l} \left(\sum_{\frac{d\gamma_k}{m} \le j < \frac{d\gamma_{k+1}}{m}} n_j \right) \log|\varphi_k'(0)| \right) \\
&\le \exp\left(\sup_{\mathbf{T}^d} \{\log|W(\mathbf{z})F(\Phi(\mathbf{z}))|\} - \sum_{k=0}^{l} \left(\sum_{\frac{d\gamma_k}{m} \le j < \frac{d\gamma_{k+1}}{m}} n_j \right) \log|\varphi_k'(0)| + O_h(1) \right).
\end{aligned} \tag{6.5.12}$$

Here, since $|\varphi_l'(0)| > 1$ and this is an upper bound, we have legitimately removed the dash proviso in the inner summation over j. We further rework (6.5.12) using (6.4.3) and apply an Abel summation to obtain (recalling for the boundary terms that $\gamma_{l+1} = m$ and $\gamma_0 = 0$), the following bound on $\log|\beta|$:

$$\begin{aligned}
&\le \sup_{\mathbf{T}^d} \{\log|W(\mathbf{z})F(\Phi(\mathbf{z}))|\} - \frac{\alpha}{m^2} \sum_{k=0}^{l} (\gamma_{k+1}^2 - \gamma_k^2) \log|\varphi_k'(0)| + O(\epsilon\alpha) \\
&= \sup_{\mathbf{T}^d} \{\log|W(\mathbf{z})F(\Phi(\mathbf{z}))|\} - \alpha \log|\varphi_l'(0)| + \frac{\alpha}{m^2} \sum_{k=1}^{l} \gamma_k^2 \log \frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|} + O(\epsilon\alpha).
\end{aligned} \tag{6.5.13}$$

At this point we follow [CDT21, § 2.5.14] to upper-estimate the supremum term in (6.5.13). From (6.2.8), the triangle inequality yields as a pointwise upper bound over $\mathbf{z} \in \mathbf{T}^d$:

$$\log|F(\Phi(\mathbf{z}))| \le \max_{\mathbf{k}/D \in P_\epsilon^d \cap \mathbf{Z}^d} \left\{ \sum_{j=1}^{d} k_j \log|\Phi_j(z_j)| \right\} + O(\epsilon\alpha) + o(\alpha), \tag{6.5.14}$$

where the splicing notation for the univariate components of the multivariable map (which we defined in § 6.4 above)

$$\Phi(\mathbf{z}) =: (\Phi_1(z_1), \ldots, \Phi_d(z_d))$$

uses $\Phi_j(z_j) := \varphi_k(z_j)$ for the unique $k = k(j) \in \{0, \ldots, l\}$ determined by the $\boldsymbol{\gamma}$ rule spelled out in § 6.4.

6.5.15. *The numerical integration.* Upon infinitesimally scaling down $z \mapsto (1-\varepsilon)z$ the coordinate of the unit disc $\mathbf{D}$, and taking the $\varepsilon \to 0$ limit at the very end, we may and do assume that none of the holomorphic functions $\varphi_0, \ldots, \varphi_l \in \mathcal{O}(\overline{\mathbf{D}})$ have any zeros lying on the unit circle $\mathbf{T}$.

For ease of notation later, we define $T^d_{\boldsymbol{\gamma},\epsilon} := \{\mathbf{z} \in \mathbf{T}^d : \forall k, D(\mathbf{z}^{(k)}) < \epsilon\}$. We denote $d^{(k)} := \lceil \gamma_{k+1} d/m \rceil - \lceil \gamma_k d/m \rceil$ the length of the $\mathbf{z}^{(k)}$ variable block, and we write $\mathbf{z} =: e^{2\pi i \mathbf{s}}$ in block form $\mathbf{z} = (\mathbf{z}^{(0)}, \ldots, \mathbf{z}^{(l)})$, so that $\mathbf{z} \in T^d_{\boldsymbol{\gamma},\epsilon}$ is tantamount to having $D(\mathbf{s}^{(k)}) < \epsilon$ for every $k = 0, \ldots, l$. Given a vector $\mathbf{w} = (w_1, \ldots, w_d) \in \mathbf{R}^d$, let us admit a slight abuse of notation and denote by $\mathbf{w}^* =: (w_1^*, \ldots, w_d^*)$ the *increasing*[23] *rearrangement vector* of the component set of $\mathbf{w}$. By Corollary 6.2.11, the running condition $\mathbf{k}/D \in P^d_\epsilon$ implies

$$k_j^* = D(j/d) + O(\epsilon D) = (2j/d^2)\frac{\alpha}{m} + O\left(\epsilon\alpha/d\right), \qquad j = 1, \ldots, d. \tag{6.5.16}$$

Similarly, for $\mathbf{s}^{(k)} \in [0,1)^{d^{(k)}}$ with $D(\mathbf{s}^{(k)}) < \epsilon$, the increasing rearrangement $(\mathbf{s}^{(k)}))^*$ has components

$$(s^{(k)})^*_\ell = \ell/d^{(k)} + O(\epsilon), \qquad \ell = 1, \ldots, d^{(k)}. \tag{6.5.17}$$

Since all the coordinate functions $\log|\Phi_j| = \log|\varphi_{k(j)}| : \mathbf{T} \to \mathbf{R}$ are of bounded variation, Koksma's inequality (see, for instance, [CDT21, § 2.5.1] for a discussion and further references) implies that for $j \in [\lceil \gamma_k d/m \rceil, \lceil \gamma_{k+1} d/m \rceil)$,

$$\log|\Phi_j(e^{2\pi i (s^{(k)})^*_\ell})| = \log|\varphi_k(e^{2\pi i \ell/d^{(k)}})| + O(\epsilon).$$

Thus, by Koksma's inequality again, we arrive at the definition of the function (6.0.8):

$$g_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(j/d) = \log\left|\varphi_k\left(e^{2\pi i \left(j - \sum_{h=0}^{k-1} d^{(h)}\right)/d^{(k)}}\right)\right| + O(1/d).$$

The increasing rearrangement notation then reads:

$$g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(j/d) = \left(\log\left|\varphi_{k(j)}\left(e^{2\pi i \left(j - \sum_{h=0}^{k(j)-1} d^{(h)}\right)/d^{(k(j))}}\right)\right|\right)^*_j + O(1/d), \tag{6.5.18}$$

where the index $k = k(j) \in \{0, \ldots, l\}$ is determined by the rule of § 6.4, which at these arguments reads: $j/d \in [\gamma_k/m, \gamma_{k+1}/m)$.

In summary, we have proved that

$$(\Phi_j(z_j))^*_j = g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(j/d) + O(\epsilon) + O(1/d).$$

Koksma's inequality and the rearrangement inequality now yield a numerical integration estimate:

$$\begin{gathered}\mathbf{t} \in P^d_\epsilon,\ t_1 \leq \cdots \leq t_d, \quad \mathbf{z} \in T^d_{\boldsymbol{\gamma},\epsilon} < \epsilon \Longrightarrow \\ \sum_{j=1}^d 2t_j \cdot \log|\Phi_j(z_j)| \leq \sum_{j=1}^d (2j/d) g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(j/d) + O(\epsilon d) + O(1).\end{gathered} \tag{6.5.19}$$

Therefore (6.5.14) and (6.5.19) imply the following upper estimate:

$$\sup_{\mathbf{z} \in T^d_{\boldsymbol{\gamma},\epsilon}} \{\log|F(\Phi(\mathbf{z}))|\} \leq \frac{1}{d}\left(\sum_{j=1}^d 2(j/d) \cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(j/d)\right)\frac{\alpha}{m} + O(\epsilon\alpha) + O\left(\frac{\alpha}{d}\right) + o(\alpha). \tag{6.5.20}$$

[23] Or rather, nondecreasing.

At this point, Koksma's inequality applies yet again to prove the following estimate, uniformly on the well-distributed part $T^d_{\boldsymbol{\gamma},\epsilon} \subset \mathbf{T}^d$:

$$\sup_{\mathbf{z}\in T^d_{\boldsymbol{\gamma},\epsilon}} \{\log|F(\Phi(\mathbf{z}))|\} \leq \frac{\alpha}{m}\cdot\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt + O(\epsilon\alpha) + O\left(\frac{\alpha}{d}\right) + o(\alpha). \quad (6.5.21)$$

6.5.22. *Noise canceling.* To handle the complementary (badly distributed) part of the integration torus $\mathbf{T}^d$, we select the "sufficiently big" exponent M of the damping Vandermondian:

$$M := \left\lfloor \frac{\sup_{\mathbf{T}}\log|\Phi|}{c'(\epsilon)}\frac{\alpha}{d^2}\right\rfloor, \quad (6.5.23)$$

where $c(\epsilon)$ is the function from (6.5.9), and we recall that we have assumed the condition $d \geq d_0(\epsilon)$ in that lemma. On the poorly distributed part $\mathbf{T}^d \smallsetminus T^d_{\boldsymbol{\gamma},\epsilon}$ we get:

$$\sup_{\mathbf{z}\in\mathbf{T}^d\smallsetminus T^d_{\boldsymbol{\gamma},\epsilon}} \{\log|W(\mathbf{z})F(\Phi(\mathbf{z}))|\} = O(\epsilon\alpha) + o(\alpha). \quad (6.5.24)$$

Putting together (6.5.21), (6.5.24), and Lemma 6.5.3, we derive the uniform estimate

$$\sup_{\mathbf{z}\in\mathbf{T}} \{\log|W(\mathbf{z})F(\Phi(\mathbf{z}))|\} \leq \frac{\alpha}{m}\cdot\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt + O_\epsilon\left(\frac{\log d}{d}\alpha\right) + O(\epsilon\alpha) + o(\alpha). \quad (6.5.25)$$

6.5.26. *The Cauchy bound.* Our upper bound on the leading $\mathbf{x}^{\mathbf{n}}$ coefficient β now follows as the combination of (6.5.13) and (6.5.25)

$$\begin{aligned}\log|\beta| \leq \frac{\alpha}{m}\cdot\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt - \alpha\log|\varphi_l'(0)| + \frac{\alpha}{m^2}\sum_{k=1}^{l}\gamma_k^2\log|\varphi_k'(0)| \\ + O_\epsilon\left(\frac{\log d}{d}\alpha\right) + O(\epsilon\alpha) + o(\alpha).\end{aligned} \quad (6.5.27)$$

This asymptotic inequality, upon taking the limits in the order $\alpha\to\infty, d\to\infty$, and $\epsilon\to 0$, already proves the special case $\mathbf{b}=\mathbf{0}, \mathbf{e}=\mathbf{0}$ of the theorem, whereby $\beta\in\mathbf{Z}\smallsetminus\{0\}$ is a nonzero rational integer and therefore at least one in magnitude. To complete the general case, it remains to estimate the denominator of the leading order coefficient $\beta\in\mathbf{Q}^\times$ in $F(\mathbf{x})$.

6.6. **Denominator arithmetic.** This is a new aspect which we did not encounter in [CDT21]. We consider all the possible combinations (6.2.8) with $c_{\mathbf{i},\mathbf{k}}\in\mathbf{Z}$, and in those, we estimate prime-by-prime the worst possible denominator that may arise in a leading order monomial coefficient β, under the premises of Lemma 6.2.6 and the denominator types (6.0.9). We prove $\exp\left(\alpha\tau^\flat(\mathbf{b}) + o(\alpha)\right)$ as the best-possible (exact) formula in the $\mathbf{e}=\mathbf{0}$ case. For the general case with added integrals, the exact denominator worst-case analysis seems subtle — especially if in the actual $m=14$ case in § 13 of our main application one tries to consider the finer denominators we indicate by Remark 10.2.3; — but we provide a handy upper estimate which turns out to be the quantity $\exp\left(\alpha\tau^\flat(\mathbf{b}) + \alpha\tau^\sharp(\mathbf{e}) + o(\alpha)\right) = \exp\left(\alpha\tau(\mathbf{b};\mathbf{e}) + o(\alpha)\right)$ of the statement of Theorem 6.0.2. We suspect our estimate to be pretty sharp in the case that we use for the proof of Theorem A.

6.6.1. *A preview on $\tau^\sharp$.* It is plain from the way these growth rates $\tau^\flat(\mathbf{b})$ and $\tau^\sharp(\mathbf{e})$ are added up that we are separately estimating the extra denominators that the factors n^{e_i} introduce from (6.0.9). The parameter $\xi \in [0, m]$ of the definition (6.0.6) of $\tau^{\#}(\mathbf{e})$ is used for the cutoff in Lemma 5.0.4 to decide for which primes p to estimate the added power of p in $\operatorname{den}(\beta)$ based on the lemma, and for which primes to estimate it based, instead, directly on the remark that every product $f_{i_1}(x_1)\cdots f_{i_d}(x_d)$ has only a limited number of factors involving "extra n denominators" in (6.0.9): namely, precisely $(\sum_{i=1}^m e_i)\, d/m$ of the d factors contribute, if we count with multiplicities n^{e_i}. For the primes $p \leq \xi D$, we use the latter trivial estimate; for the range $p > \xi D$, we estimate by Lemma 5.0.4 using that the leading $\prec$-order exponent vector $\mathbf{n}$ is, upon relabeling the variables, close to $(mD/d, 2mD/d, \ldots, dmD/d)$.

This means looking respectively into the left-hand side and the right-hand side of the identity $\sum_{\mathbf{i},\mathbf{k}} c_{\mathbf{i},\mathbf{k}} \mathbf{x}^{\mathbf{k}} f_{i_1}(x_1)\cdots f_{i_m}(x_m) = \sum_{\mathbf{m}} \beta_{\mathbf{m}} \mathbf{x}^{\mathbf{m}}$. In our $d \to \infty$ asymptotic, our $\prec$-leading exponent vector $\mathbf{n} \approx (mD/d, 2mD/d, \ldots, dmD/d)$ realizes once again the cross-variables dimension, with an integration variable $t := jm/d \in [0, m]$ that leads up to the function $\max(e_i) \cdot I_\xi^m(\xi)$ of Definition 6.0.1. More precisely, recalling that we will take $\epsilon \to 0$ in the end, which will automatically force $\delta \to 0$, the latter emerges as the $\int_0^{m+\delta}$ Riemann integral of the

$$\frac{\max_i(e_i)}{D} \log \frac{[\max(1, tD - D), tD]}{\gcd\{[1, \ldots, \xi D], [\max(1, tD - D), tD]\}}\, dt$$

estimates from Lemma 5.0.4. As this latter estimate goes "across the variables," it only "sees" the n^{e_i} exponents through their common capping $\max_i(e_i)$, based on the remark that, separately in every variable x_j, all the terms $[x_j^n]\, f_i(x_j)$ have added denominators multiplying at most by $n^{\max_i(e_i)}$; this has to be taken uniformly in $i \in \{1, \ldots, m\}$ (hence the maximum over i), as each given function species f_i will occur from some product of (6.2.8) at every single coordinate x_j. To leverage that many of the f_i could have smaller added denominators n^{e_i} than the common $n^{\max_i(e_i)}$ capping of this cross-variable denominator estimation, we use a balancing parameter ξ and directly estimate the primes $p \leq \xi D$ from the multiplicity density $(\sum_{i=1}^m e_i)\, d/m$ of affected factors in each product $f_{i_1}(x_1)\cdots f_{i_m}(x_d)$, in which an extra p^{e_i} denominator could possibly be hiding (this is a conservative estimate!).

We now execute both points $\tau^\flat(\mathbf{b})$ and $\tau^{\#}(\mathbf{e})$ of this $\operatorname{den}(\beta)$ majorization program. In these denominator estimates, the essential point is that our $\prec$-minimal exponent $\mathbf{n} \in (m+\delta)D \cdot P_\epsilon^d$ in $F(\mathbf{x})$ has uniformly distributed components, but the corresponding information on $\mathbf{k} \in D \cdot P_\epsilon^d$ is now ignored. (It is conceivable that the latter could be also exploited to give a more precise bound; however, we were unable to do that in our applications at hand.)

This is the exact opposite to the archimedean growth estimate in § 6.5.

6.6.2. *The $\tau^\flat(\mathbf{b})$ piece.* Consider any of the lowest order exponent vectors

$$\mathbf{n} = (n_1, \ldots, n_d) \in (m+\delta)D \cdot P_\epsilon^d,$$

as given by Lemma 6.2.6. Recall that in § 6.4 we relabeled the coordinates to assume — simply for a notational convenience — that our $\mathbf{n}$ has nondecreasing components: $n_1 \leq \cdots \leq n_d$. Then, by Corollary 6.2.11, we have

$$\begin{aligned} n_j &\leq (1+\epsilon) mD(j/d) + 2m\epsilon D \\ &\leq mD(j/d) + 3m\epsilon D, \qquad j = 1, \ldots, d. \end{aligned} \tag{6.6.3}$$

We need to compute the lowest common denominator of all the nonzero rational numbers $\beta \in \mathbf{Q}^{\times}$ that may arise as the $\mathbf{x}^{\mathbf{n}}$ coefficient in any product

$$x_1^{k_1}\cdots x_d^{k_d}\cdot g_{\pi(1)}(x_1)\cdots g_{\pi(d)}(x_d) \in \mathbf{Q}[\![\mathbf{x}]\!] \tag{6.6.4}$$

with some arbitrary $\mathbf{k} \in D \cdot P_{\epsilon}^{d}$, some arbitrary permutation π of $\{1,\dots,d\}$, and for each $i \in \{1,\dots,m\}$, some arbitrary formal functions

$$g_{1+(i-1)d/m}(x),\dots,g_{id/m}(x) \in \bigoplus_{n=0}^{\infty} \frac{x^n}{[1,\dots,b_{i,1}\cdot n]\cdots[1,\dots,b_{i,r}\cdot n]}\,\mathbf{Z}. \tag{6.6.5}$$

This means nothing more nor less than the lowest common multiple of all products

$$\prod_{i=1}^{m}\prod_{s=1}^{d/m}\prod_{h=1}^{r}\left[1,\dots,b_{i,h}\cdot n_{\pi((i-1)d/m+s)}\right], \qquad \pi \in S_d, \tag{6.6.6}$$

as π ranges over all permutations of $\{1,\dots,d\}$.

We handle (6.6.6) with a prime-by-prime determination of the maximizing valuation. The following simple lemma is where the special condition (6.0.3) on the denominators shape matrix $\mathbf{b}$ is used, in all our theorems in §§ 6, 7.

Lemma 6.6.7. *For every prime p, every vector $(c_1,\dots,c_m) \in \mathbf{N}^m$ of the form*

$$0 = c_1 = \cdots = c_u < c_{u+1} = \cdots = c_m =: c, \tag{6.6.8}$$

and every nondecreasing sequence $n(1) \le \cdots \le n(km)$ consisting of km positive integers, the following equality holds:

$$\begin{aligned}\max_{\pi\in S_{km}} \quad &\mathrm{val}_p\left\{\prod_{i=1}^{m}\prod_{s=1}^{k}\left[1,\dots,c_i\cdot n\left(\pi\left((i-1)k+s\right)\right)\right]\right\}\\ &= \mathrm{val}_p\left\{\prod_{i=1}^{m}\prod_{s=1}^{k}\left[1,\dots,c_i\cdot n\left((i-1)k+s\right)\right]\right\}.\end{aligned} \tag{6.6.9}$$

In other words, as $\pi \in S_{km}$ ranges through all permutations of $\{1,\dots,km\}$, the identity permutation $\pi = \mathrm{id}$ maximizes the p-adic valuation in (6.6.9).

Proof. The condition (6.6.8) simplifies the requisite product (6.6.9) to

$$\prod_{i=u+1}^{m}\prod_{s=1}^{k}\left[1,\dots,c\cdot n\left(\pi\left((i-1)k+s\right)\right)\right]. \tag{6.6.10}$$

The lowest common multiple $[1,\dots,N]$ of the first N positive integers has p-adic valuation equal to $\left\lfloor \frac{\log N}{\log p}\right\rfloor$, and so the quantity in (6.6.9) under maximization is exactly equal to

$$\sum_{i=u+1}^{m}\sum_{s=1}^{k}\left\lfloor \frac{\log c}{\log p} + \frac{\log n\left(\pi\left((i-1)k+s\right)\right)}{\log p}\right\rfloor. \tag{6.6.11}$$

From the km positive integers $\{n(1),\dots,n(km)\}$, we have to pick $k(m-u)$ with pairwise distinct indices to maximize the sum (6.6.11). Clearly this is maximized by picking the $k(m-u)$ largest available numbers $n(\bullet)$, so in particular our monotonicity assumption on $n(\bullet)$ gives that (6.6.11) is maximized by the identity permutation $\pi = \mathrm{id}$, with maximum

$$\sum_{j=ku+1}^{km} \Big\lfloor \frac{\log c}{\log p} + \frac{\log n(j)}{\log p} \Big\rfloor. \qquad \square$$

Applying (6.6.3) on the nondecreasing sequence $\mathbf{n}$, together with our condition (6.0.3) on the $m \times r$ array, we find by Lemma 6.6.7 that as soon as $D \gg_\epsilon 1$, all the lowest common multiple products (6.6.6) divide

$$\prod_{i=1}^{m} \prod_{s=1}^{d/m} \prod_{h=1}^{r} \left[1, \dots, b_{i,h} \cdot \big((i-1)D + smD/d\big) + \epsilon BD\right], \tag{6.6.12}$$

where the constant $B := m \cdot \max_{i,h}\{b_{i,h}\}$.

By the prime number theorem, the lowest common multiple cap (6.6.12) evaluates in the $D \to \infty$ asymptotic to

$$\begin{aligned} &\exp\left(\sum_{i=1}^{m}\sum_{s=1}^{d/m}\sum_{h=1}^{r} \big(b_{i,h} \cdot \big((i-1)D + smD/d\big) + O(\epsilon D)\big)\right) \\ &= \exp\left(mD \sum_{i=1}^{m}\sum_{s=1}^{d/m} \sigma_i \cdot \big((i-1)/m + s/d\big) + O(\epsilon dD)\right) \\ &= \exp\left(\alpha \sum_{i=1}^{m} \sigma_i \int_{(i-1)/m}^{i/m} 2t\,dt + O(\epsilon\alpha) + o_{d\to\infty}(\alpha)\right) \\ &= \exp\Big(\alpha\tau^\flat(\mathbf{b}) + O(\epsilon\alpha) + o_{d\to\infty,\epsilon\to 0}(\alpha)\Big), \end{aligned} \tag{6.6.13}$$

recalling our definition (6.2.10) of the vanishing order parameter $\alpha = mdD/2$.

Remark 6.6.14. The statement of Lemma 6.6.7 ceases to be true if the condition (6.6.8) is relaxed to an arbitrary monotonic $0 \le c_1 \le \cdots \le c_m$. Thus, with $\tau^\flat(\mathbf{b}) = \frac{1}{m^2}\sum_{i=1}^m (2i-1)\sigma_i$ as the definition in (6.0.4), the proof of the theorem would no longer hold if we relaxed the crude capping (6.0.3) of our denominator types to an arbitrary matrix $\mathbf{b}$ having columns with nondecreasing components. $\triangle$

Remark 6.6.15. Unlike for the archimedean growth estimate in § 6.5.15, our computation here ignored the uniform distribution constraint $\mathbf{k} \in D \cdot P_\epsilon^d$ inside the trivial estimate $\mathbf{k} \in [0, D]^d$. This was how the growth rate $\tau^\flat$ was defined, not to take account of the distribution of the exponents $\mathbf{k}$ of the auxiliary polynomials; for this definition, it is an exact computation. In contrast, it was crucial for the horizontal integration idea to exploit the uniformly distributed components of the $\prec$-leading $\mathbf{x} = \mathbf{0}$ exponent $\mathbf{n} \in (m+\delta)D \cdot P_\epsilon^d$.

In principle (or in practice), upon calculating a denominator rate still more involved than the term $\mathrm{Den}_N(\mathbf{b}, \epsilon, d, \varepsilon)$ in Theorem 8.0.1, one could formulate a version of Theorem 6.0.2 in which $\tau(\mathbf{b}; \mathbf{e})$ is formally refined to a complicated limiting formula that does also takes account of the uniform components restriction on the exponent vectors $\mathbf{k} \in [0, D]^d$ in the make-up of the auxiliary function (6.2.8); and where denominator shapes still finer than our template form (7.0.1) could be considered. The denominators in Remark 10.2.3, and similar forms involving products of binomial coefficients or products of primes from restricted intervals, are the typical example to have in mind for prospective applications; for deeper studies and more complicated examples, cf. [Vio04, RV96, Sor16, Zud14, DZ14]. In the situation of our application to our main Theorem A, we did not succeed in making any (non-negligible) use of the uniformly distributed $\mathbf{k}$ for handling the more restricted denominator types of Remark 10.2.3. △

6.6.16. *The* $\tau^{\#}(\mathbf{e})$ *piece.* To estimate the denominator surplus from the extra integration denominators $n^{\mathbf{e}}$, we will separately (as an upper bound) multiply the principal denominators cap (6.6.12) by the lowest common denominator of all the possible $\mathbf{x}^{\mathbf{n}}$ coefficients $\beta \in \mathbf{Q}^{\times}$ of all possible products

$$x_1^{k_1} \cdots x_d^{k_d} \cdot g_{\pi(1)}(x_1) \cdots g_{\pi(d)}(x_d) \in \mathbf{Q}[\![\mathbf{x}]\!],$$

across all possible $\mathbf{k} \in D \cdot P_{\epsilon}^d$, some arbitrary permutation $\pi \in S_d$, and, for each $i \in \{1, \dots, m\}$, arbitrary formal functions

$$g_{1+(i-1)d/m}(x), \dots, g_{id/m}(x) \in \bigoplus_{n=0}^{\infty} \frac{x^n}{n^{e_i}} \mathbf{Z}. \tag{6.6.17}$$

This multiplies separate local estimations of the highest possible power of a denominator at every prime p. Consider $\xi \in [0, m]$ the parameter of the definition (6.0.5). We estimate differently the cases $p \leq \xi D$ and $p > \xi D$. Firstly we collect two basic standard facts, the first of which is a version of the prime number theorem, and the second, an immediate consequence:

(a) *The product of the primes* $p \leq n$ *is asymptotic to* $\exp(n + o(n))$.

(b) *The product of the proper prime powers* $p^a \leq n, a \geq 2$, *is bounded by* $\exp\left(O(\sqrt{n}) = \exp(o(n))\right.$.

Together, they imply:

(c) *The lowest common multiple* $[1, \dots, n] = \exp(n + o(n))$.

These properties prove that for the $\mathbf{x}^{\mathbf{n}}$ coefficient denominator of the "generic"

$$[1, \dots, \xi D]^{\lfloor d(\sum_{i=1}^m e_i)/m \rfloor} \cdot x_1^{k_1} \cdots x_d^{k_d} \cdot g_{\pi(1)}(x_1) \cdots g_{\pi(d)}(x_d) \in \mathbf{Q}[\![\mathbf{x}]\!], \tag{6.6.18}$$

we have:

(i) the totality of the primes $p \leq \xi D$ add only a negligible $\exp(o(\xi D)) = \exp(o(\alpha))$ factor to the denominators of (6.6.18);
(ii) the clearing factor

$$\begin{aligned}[1, \dots, \xi D]^{\lfloor d(\sum_{i=1}^m e_i)/m \rfloor} &= \exp\left(\xi d D \left(\sum_{i=1}^m e_i\right) / m + o(\alpha)\right) \\ &= \exp\left(\xi \left(\sum_{i=1}^m e_i\right) \cdot 2\alpha/m^2 + o(\alpha)\right).\end{aligned}$$

It is clear then that, up to an $\exp(o_{d\to\infty, \epsilon\to 0}(\alpha))$ factor, the lowest common denominator of all the $\mathbf{x}^{\mathbf{n}}$ coefficients of all the formal expressions (6.6.18), as the $h_j(x)$

range over (6.6.17), $\mathbf{n}$ ranges over $(m+\delta)D\cdot P_\epsilon^d$, and $\mathbf{k}$ ranges over $D\cdot P_\epsilon^d$, is a divisor of the lowest common denominator of all formal expressions

$$\frac{[1,\ldots,\xi D]^\infty}{(n_1-k_1)^{\max_i(e_i)}\cdots(n_d-k_d)^{\max_i(e_i)}},\quad \mathbf{n}\in(m+\delta)D\cdot P_\epsilon^d,\quad \mathbf{k}\in D\cdot P_\epsilon^d;$$

and that, again up to an $\exp(o_{d\to\infty,\epsilon\to 0}(\alpha))$ factor, this is a divisor[24] of

$$\prod_{j=1}^d \prod_{\substack{\text{primes } p>\xi D:\\ p \text{ divides some}\\ \text{positive integer in}\\ [mD(j/d)-D,mD(j/d)]}} p^{\max_i(e_i)}. \tag{6.6.19}$$

By Lemma 5.0.4, if $m(j/d)>1$, the inner product in (6.6.19) is asymptotic to the exponential of

$$\begin{aligned}&\left(\max_i e_i\right)D\left(\sum_{h=1}^{\lfloor (m(j/d)-1)/\max(1,\xi)\rfloor} 1/h\right)\\ &+\left(\max_i e_i\right)D\max\left\{\frac{m(j/d)}{\lfloor (m(j/d)+\max(0,\xi-1))/\max(1,\xi)\rfloor}-\xi,0\right\}+o(D).\end{aligned}$$

In the case $m(j/d)\le 1$, the inner product in (6.6.19) is asymptotic to the exponential of

$$\left(\max_i e_i\right)D\max\{0,m(j/d)-\xi\}+o(D).$$

Hence, recollecting our Definition 6.0.1 of the integrated LCM cost function $I_u^v(w)$, we find that as $d\to\infty$, so that the discrete variable $t:=m(j/d)$ converges to the continuous Lebesgue measure of the segment $[0,m]$, the horizontal integration computes the asymptotic full denominator product (6.6.19) to the following, up to an $\exp(o_{d\to\infty,\epsilon\to 0}(\alpha))$ factor:

$$\begin{aligned}&\exp\left((dD/m)\left(\max_i e_i\right)\cdot I_\xi^m(\xi)+o_{d\to\infty,\epsilon\to0}(dD)\right)\\ &=\exp\left((2\alpha/m^2)\left(\max_i e_i\right)\cdot I_\xi^m(\xi)+o_{d\to\infty,\epsilon\to0}(\alpha)\right).\end{aligned} \tag{6.6.20}$$

All in all, we obtain for any $\xi\in[0,m]$ the upper estimate

$$\exp\left((2\alpha/m^2)\cdot\left(\xi\sum_{i=1}^m e_i+\left(\max_{1\le i\le m} e_i\right)\cdot I_\xi^m(\xi)\right)+o(\alpha)+o_{d\to\infty,\epsilon\to0}(\alpha)\right) \tag{6.6.21}$$

on the addition to the denominator from the $n^{\mathbf{e}}$ factors in (6.0.9) in any leading order coefficient of our auxiliary function $F(\mathbf{x})$ in Lemma 6.2.6.

Our definition of the rate $\exp\left(\alpha\cdot\tau^\sharp(\mathbf{e})+o(\alpha)\right)$ *is as the minimum of the total added denominators estimate* (6.6.21) *over the parameter* $\xi\in[0,m]$.

[24] Even upon including all $\mathbf{k}\in[0,D]^d$, that is once again ignoring the $\mathbf{k}\in P_\epsilon^d$ constraint.

6.7. **Proof of Theorem 6.0.2.** We combine the upper bound (6.5.27) on the leading $\mathbf{x^n}$ coefficient $\beta \in \mathbf{Q}^{\times}$ with the added up upper estimates that we computed in § 6.6.2 and § 6.6.16 on the denominator $\mathrm{den}(\beta) \in \mathbf{N}_{>0}$ of that coefficient. The latter give:

$$\log|\beta| \geq -\alpha \cdot \Big(\tau^{\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e})\Big) + o_{d\to\infty,\epsilon\to 0}(\alpha). \tag{6.7.1}$$

The former simplifies to:

$$\begin{aligned}\log|\beta| \leq \frac{\alpha}{m}\cdot \int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt - \alpha\log|\varphi_l'(0)| \\ + \frac{\alpha}{m^2}\sum_{k=1}^{l}\gamma_k^2\log\frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|} + o_{d\to\infty,\epsilon\to 0}(\alpha).\end{aligned} \tag{6.7.2}$$

The combination of (6.7.1) and (6.7.2) sifts out in the $\alpha\to\infty, d\to\infty, \epsilon\to 0$ limit to

$$m \leq \frac{\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt + \frac{1}{m}\sum_{k=1}^{l}\gamma_k^2\log\frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|}}{\log|\varphi_l'(0)| - \tau^{\flat}(\mathbf{b}) - \tau^{\sharp}(\mathbf{e})}, \tag{6.7.3}$$

which is precisely our claimed holonomy bound. □

At this point, a reader primarily interested in the proof of Theorems A and C can skip directly ahead to § 9 on a first reading.

6.8. **Completion of the proof of Theorem 2.8.4.** In § 2.11, by a direct application of Theorem 2.5.1, we already proved the property $(*)$ in Theorem 2.8.4 towards the arithmetic characterization of the $\log^2(1-x)$ function. This was the case — the one of relevance to the sample application to $\mathbf{Q}$-linear independence proofs that we gave in § 2.11.12 — that the minimal order differential operator $\mathcal{L}$ has an essential singularity at the "fourth puncture" $x=\delta$ from the statement of the theorem. With the feature of the multiple maps $\boldsymbol{\varphi}$ in Theorem 6.0.2, we can now complete the proof of the full Theorem 2.8.4 by handling the case that $x=\delta$ is at most an apparent singularity of $\mathcal{L}$.

Proof of Theorem 2.8.4. From the discussion in § 2.11, as the setup of the theorem-under-proof implies that our type $[1,\dots,n][1,\dots,n/2]$ formal function $f(x)\in\mathbf{Q}[\![x]\!]$ has a meromorphic pullback under $\varphi(z) := \frac{8(z+z^3)}{(1+z)^4}$, where already $|\varphi'(0)| = 8 > \tau = 3/2$, we certainly get the existence of a minimal-order nonzero linear differential operator $\mathcal{L}$ over $\mathbf{Q}(x)$ satisfying $\mathcal{L}(f)=0$. It remained to cover the case that the linear ODE $\mathcal{L}(F)=0$ has a full set of *meromorphic* solutions in a neighborhood of $x=\delta$. Upon multiplying by a nonzero polynomial $Q\in\mathbf{C}[x]\smallsetminus\{0\}$ to clear up the possible meromorphic poles, this is equivalent to the assumption that the holomorphic function germ $Q(x)f(x)\in\mathbf{C}[\![x]\!]$ is analytically continuable as a *holomorphic* function along all paths in $\mathbf{P}^1\smallsetminus\{0,1,\infty\}$. By Proposition 2.9.3, this condition in turn furnishes a meromorphic pullback $\varphi^*f = (\varphi^*(Qf))/\varphi^*Q \in \mathcal{M}(\overline{\mathbf{D}})$ under all holomorphic mappings $\varphi:\overline{\mathbf{D}}\to\mathbf{C}\smallsetminus\{1\}$ subject to $\varphi^{-1}(0)=\{0\}$.

The reason that the previous argument breaks down in this case is that, in the absence of the fourth singularity $\delta\notin\{0,1,2,\infty\}$, there is no longer a reason for the $\mathbf{Q}(x)$-linear independence of $f(x/(x-1))$ from the four other functions in (2.11.10), and we only have $m=4$ with the functions

$$f_1(x) := 1,\quad f_2(x) := \log(1-x),\quad f_3(x) := \log^2(1-x),\quad f_4(x) := f(x), \tag{6.8.1}$$

and the choice of denominators type given by $\frac{1}{2}\mathbf{b}_0$ and $\mathbf{e} = \mathbf{0}$ of Example 6.0.20. On the other hand, the absence of δ spares us the trouble to have the analytic mapping φ to necessarily cover the value δ only once, and we can take a completely arbitrary $\varphi : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \{1\}$ subject only to $\varphi^{-1}(0) = \{0\}$.

Suppose for the contradiction that there exists a fourth $\mathbf{Q}(x)$-linearly independent function $f(x)$ in (6.8.1) still of the type $[1, \dots, n/2][1, \dots, n]$, and therefore completing the combined type $\frac{1}{2}\mathbf{b}_0$ from Example 6.0.20. With $S = T = \emptyset$ and $\mathcal{H}$ taken as the four-dimensional $\mathbf{Q}(x)$-linear span of (6.8.1), we use the symmetrization dictionary $\varphi_{Y(2)} \rightsquigarrow \varphi_{Y_0(2)}$ described in Basic Remark 9.0.20 in the section § 9 on the $y := x + x/(x-1) = x^2/(x-1)$ descent, with technical details given by Lemma 9.0.3 (on the algebraic and arithmetic sides) and Lemma 9.0.13 (on the analytic side). Explicitly, using the involution $w(x) := x/(x-1)$ and the symmetrization coordinate $y := x + w(x) = xw(x) = x^2/(x-1)$, we have for $\mathcal{H}^{w=1}$ the four-dimensional $\mathbf{Q}(y)$-vector space of the denominator type denoted $\mathbf{b}_0$ in Example 6.0.20, with $\mathbf{e} = \mathbf{0}$ (no added integrations), and spanned by 1, $\sqrt{y(4-y)} \arcsin \frac{\sqrt{y}}{2}$, $\left(\arcsin \frac{\sqrt{y}}{2}\right)^2$, and the symmetrizations of $f(x)$. In the notation of Basic Remark 9.0.20, where in particular $h = \lambda^2/(\lambda - 1) = -256q + \cdots$ denotes a hauptmodul (9.0.1) of $Y_0(2)$ written out in the coordinate $q := e^{2\pi i \tau}$, we select for our ambiance the analytic mapping

$$\varphi_{Y_0(2)} := h \circ \mathrm{Gob}(1/2, 10, 3) \in \mathcal{O}(\overline{\mathbf{D}}), \tag{6.8.2}$$

where $\mathrm{Gob}(1/2, 10, 3) : \overline{\mathbf{D}} \hookrightarrow \mathbf{D}$ is the domain described in § A.2. Lemma A.2.2 computes $|\mathrm{Gob}'(1/2, 10, 3)(0)| = 198/505$ for the conformal radius of this domain, giving

$$|\varphi'_{Y_0(2)}(0)| = 256 \cdot \frac{198}{505} = \exp(4.608886\ldots) \tag{6.8.3}$$

for the conformal size of our ambient analytic map. It is usefully large in comparison to the denominators growth rate $\tau(\mathbf{b}_0) = 21/8 = 2.625$ from Example 6.0.20.

Here Corollary 9.0.19, as interpreted by Basic Remark 9.0.20 and once again using a suitable polynomial multiplier $Q \in \mathbf{C}[y] \smallsetminus \{0\}$ to clear all the possible meromorphic poles, proves that our choice (6.8.2) resolves analytically the four-dimensional holonomic $\mathbf{Q}(y)$-module $\mathcal{H}^{w=1}$:

$$\dim_{\mathbf{Q}(y)} \mathcal{H}^{w=1} = 4, \qquad \varphi^*_{Y_0(2)} \mathcal{H}^{w=1} \subset \mathcal{M}(\overline{\mathbf{D}}). \tag{6.8.4}$$

This is all conditional on the supposed falsity of the theorem under proof. It is to this $\mathbf{Q}(y)$ situation that we apply Theorem 6.0.2, with $m := 4$ and the following choices of the intermediate maps φ_k and division parameters γ_k:

$$\begin{aligned} l := 2; \gamma_1 := 3/5, \gamma_2 := 2; \\ \varphi_0(z) := \varphi_{Y_0(2)}\left(e^{-5} z\right), \\ \varphi_1(z) := \varphi_{Y_0(2)}\left(e^{-1/2} z\right), \\ \varphi_2(z) := \varphi_{Y_0(2)}(z). \end{aligned} \tag{6.8.5}$$

`Mathematica` then yields a holonomy quotient (6.0.10) of slightly under < 3.9, which is the desired contradiction to complete the proof of the $\mathbf{Q}(x)$-linear independence of the four original functions (6.8.1).

Finally, the integral refinement over $\mathbf{Q}\left[x, \frac{1}{x}, \frac{1}{1-x}\right]$ follows at this point from the Hermite–Lindemann–Weierstrass and Mahler theorems in transcendence, exactly as in the proof of Theorem 2.7.2; and then the upgrade from $\mathbf{Q}\left[x, \frac{1}{x}, \frac{1}{1-x}\right]$ to $\mathbf{Q}\left[x, \frac{1}{1-x}\right]$ follows from our strict $[1,\dots,n][1,\dots,n/2]$ denominators requirement exactly as in point (ii) in the proof of Theorem 2.7.2. □

Example 6.8.6. With $m = 3$ and the true functions

$$1, \sqrt{y(4-y)}\,\arcsin\frac{\sqrt{y}}{2}, \left(\arcsin\frac{\sqrt{y}}{2}\right)^2$$

now using the type

$$\mathbf{b} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \\ 1 & 2 \end{pmatrix}, \qquad \mathbf{e} = (0,0,0), \qquad \tau(\mathbf{b};\mathbf{e}) = \frac{7}{3},$$

a short numerical experimentation, which we have not attempted to make rigorous, suggests that for $l = 2$ (two division points) and maps of the form

$$\varphi_{Y_0(2)}(z) := h\left(\mathrm{Gob}(r,e,f)(\delta z)\right), \quad \varphi_k(z) := \varphi_{Y_0(2)}\left(\gamma_k z\right), \quad k = 0,1,2,$$

the minimizing holonomy bound on the three functions should probably be attained at (for example) about the choice

$$\begin{aligned} &(r,e,f) \approx (0.55, \infty, 5); \quad (\gamma_1,\gamma_2) \approx (0.19, 0.65), \\ &(r_0, r_1) \approx \left(e^{-4.3}, e^{-0.76}\right), \quad \delta \approx 0.77, \end{aligned}$$

with holonomy quotient value (6.0.10) being at ≈ 3.239. △

7. Convexity in Bost's slopes method

We begin with the following clean refinement of Theorem 2.5.1, which finally we prove in this section by a single variable method based on [BC22, § 5]. This simple result by itself suffices for our application to Theorems A and C, although (in the case of the former) only by the narrowest of margins. The tenor of this section, whose main results are stated in § 7.1 after a short introduction, is what we are calling the *convexity input* that leads up to sharpened holonomy bounds. The improvements are usually fairly small, but they are significant enough to comfortably pass the numerical margin in the requisite numeric in the proof of Theorem A, and thereby make fully convincing the Arakelov theory path to the irrationality proof of $L(2,\chi_{-3})$.

Theorem 7.0.1. *With the same standing assumptions of Theorem 6.0.2, consider a holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{C}, 0)$ with derivative (conformal size) satisfying the condition*

$$\log|\varphi'(0)| > \max\{\sigma_m, \tau(\mathbf{b};\mathbf{e})\}. \tag{7.0.2}$$

Suppose there exists an m-tuple $f_1,\dots,f_m \in \mathbf{Q}[\![x]\!]$ of $\mathbf{Q}(x)$-linearly independent formal functions with denominator types of the form

$$f_i(x) = a_{i,0} + \sum_{n=1}^{\infty} a_{i,n} \frac{x^n}{n^{e_i}[1,\dots,b_{i,1}\cdot n]\cdots[1,\dots,b_{i,r}\cdot n]}, \qquad a_{i,n} \in \mathbf{Z},$$

such that $f_i(\varphi(z)) \in \mathbf{C}[\![z]\!]$ *is the germ of a meromorphic function on* $|z| < 1$ *for all* $i = 1, \dots, m$*. Then*

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \tag{7.0.3}$$

In particular, the formal functions $f_1, \dots, f_m$ *are holonomic.*

If moreover all functions f_i *are* a priori *assumed to be holonomic, the condition* (7.0.2) *can be relaxed to* $|\varphi'(0)| > e^{\tau(\mathbf{b};\mathbf{e})}$.

When $\mathbf{e} = \mathbf{0}$, then (as previously noted after the statement of Theorem 6.0.2) $\tau(\mathbf{b};\mathbf{e})$ coincides with the $\tau(\mathbf{b})$ of Theorem 2.5.1. Hence Theorem 2.5.1 is an immediate corollary of Theorem 7.0.1.

As discussed in § 2.1, the $\mathbf{b} = \mathbf{0}$, $\mathbf{e} = \mathbf{0}$ case of $\mathbf{Z}[\![x]\!]$ functions was established by Bost and Charles [BC22, Corollary 8.3.5]. Charles explained to us that their method can be modified to take denominators into account and obtain the following weaker form[25] of (7.0.3) for a starting bound:

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - (\sigma_m + \max_{1\leq i\leq m} e_i)}. \tag{7.0.4}$$

The basic idea of Theorem 7.0.1 is to use the same archimedean estimate as in [BC22], but incorporate into it a closer nonarchimedean evaluation height counterpart that, sufficiently for our purposes in the paper, improves the denominator of the initial bound (7.0.4). Remarkably, despite two seemingly rather different methods being used (the concentration of measure method exploiting multivariable approximations in § 6, and a single variable evaluation heights scheme in the present section), the final denominator term is exactly the same in both Theorems 6.0.2 and all the results in the present section. The measure concentration method is nevertheless theoretically more precise in the general denominators aspect, even though no difference is made to any case of relevance to this paper. In the next section § 8, we formulate the most precise of our holonomy bounds by blending together the measure concentration feature of § 6 with the Bost–Charles feature of the present § 7.

For the present section and the next, we use Bost's slopes inequality in Arakelov theory. A practically equivalent framework would be the elementary dynamic box principle of § B; we stick to the former choice for variety in our paper, and because the archimedean evaluation height estimate requires in any event the appeal to some relatively deep theorems in Arakelov theory. We hope that the elementary evaluation heights arrangement in § B could nevertheless be helpful to some readers as an introduction to the tenor of the more elaborate method (due to Bost) that we take up here.

We also note that one can more intrinsically formulate the proof of (7.0.3) in terms of Bost's theta invariants h^0_θ and h^1_θ as in [Bos20, BC22]. The latter pursue the concept of absolute dimension for the "space of global sections" of a (countably) infinite-dimensional Hermitian vector bundle over an arithmetic curve, under the traditional analogy between number fields and algebraic curves over finite ground fields. We do not pursue this approach here as the subsequent "convexity improvements" of the archimedean growth term in (7.0.3) seem to be more of an analytic than a geometric nature.

[25] Conditional, as always, on the positivity of the denominator.

7.1. **Improvements from convexity.** In § B.3, the dynamic pigeonholing involves a certain upper bound (B.3.2) on the interval of possibilities for each Taylor coefficient of a power series $F(x) = \sum_{i=1}^m Q_i(x) f_i(x) \in \mathbf{Q}[\![x]\!]$ in the range of the auxiliary evaluation map, given all previous coefficients of that power series. This upper bound is the result of estimating a Cauchy contour integrand (B.3.1) over $\mathbf{T}$ to express the z^n coefficient of the $x = \varphi(z)$ pullback of an element of $\psi_D\left(E_D^{(n)}\right) \subset x^n \cdot \mathbf{Q}[\![x]\!]$. As in Theorem 6.0.2, there is no particular reason to stick to a single fixed analytic map φ for each filtration layer n in this set of analytic estimates. Notably, depending on n/D, the ambient map φ, and the choice of Euclidean metric structure in the evaluation module E_D (cf. § 7.2 below), there is a certain optimal choice of an intermediate radius $r = r(n) \in (0,1]$ for estimating this *n-th archimedean evaluation height* analytically via the pullback by $x = \varphi(r(n)z)$; the only difference is that now, in the single-variable analysis, the integration procedure over n/D is "vertical" along the vanishing order, rather than "horizontal" across variables. The choice of $r(n)$ has a geometric significance with convex hulls; incidentally giving another nuance to the name *slopes method*. It aligns with the well-known fact that the Nevanlinna growth characteristic, and various cognate quantities, are convex increasing functions of the logarithm of the radius.

In this section, we compute these optimal choices $r(n)$ for two types of Euclidean metrics in the evaluation module: the *Bost–Charles metric* from [BC22], and an explicit family of *binomial metric weights* $\lambda t^r + \mu t$ depending on three real parameters (r, λ, μ), which are better amenable to numerical computation and still tend in practice to return close-to-optimal bounds for the best triple (r, λ, μ). With either of these variations, the results of this section alone (which are independent of § 4 and § 6, see also § 1.3) lead to a proof of Theorems A and C. We spell them out in § 7.1.1 and § 7.1.12 next, and prove them in § 7.4 and § 7.5 after preparations in § 7.2. Along the way, the proof of the more basic Theorem 7.0.1 appears in § 7.3. After that, in § 7.6, we discuss a further improvement that lines up with — and theoretically[26] strengthens — Theorem 6.0.2 of the previous section.

7.1.1. *The Bost–Charles characteristic.* Inspired by [BC22, § 5], we introduce a (doubled) Nevanlinna-type growth characteristic, sticking for simplicity to the case of relevance here of *holomorphic* (rather than meromorphic) disc maps $\overline{\mathbf{D}} \to \mathbf{C}$. Crucially for this approach, there turns out to be an interpretation of this growth characteristic as an arithmetic intersection number in the Bost–Charles theory.

Definition 7.1.2. For a nonconstant holomorphic[27] mapping $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$, define the *Bost–Charles characteristic function*

$$\widehat{T}(\cdot, \varphi) : (0,1] \to \mathbf{R}, \qquad \widehat{T}(r, \varphi) := \iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(rw)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w). \tag{7.1.3}$$

As with the usual Nevanlinna and Ahlfors–Shimizu characteristics, see [Nev70, § 3.3.5] or [BG06, Remark 13.3.8], we have:

Lemma 7.1.4. *The Bost–Charles characteristic $\widehat{T}$ is a convex increasing function of* $\log r$.

[26] At least in all the cases that we encountered in practice; see Remark **??**

[27] In the general meromorphic case, which we will not consider here, a suitable polar counting term would have to be added.

Proof. The Poisson–Jensen formula allows the following rewriting of the double continuous integral $\widehat{T}(r,\varphi)$ as a single continuous integral of a discrete sum:

$$\widehat{T}(r,\varphi) = \int_{\mathbf{T}} \left\{ \log|\varphi(z)| + \sum_{\substack{u\in\mathbf{D} \\ \varphi(u)=\varphi(z)}} \log^+ \frac{r}{u} \right\} \mu_{\mathrm{Haar}}(z). \tag{7.1.5}$$

We have substituted here $u := rw$. The lemma now follows upon remarking that the finite sum in the curly brackets is itself a convex increasing function of $\log r$ for each given $z \in \mathbf{T}$. ☐

We give two essentially equivalent formulations for the main theorem of this section.

Theorem 7.1.6. *Assume the same conditions and notation as in Theorem 7.0.1. Let*

$$1 = r_l > r_{l-1} > \cdots > r_0 > 0$$

be a sequence of subradii, and consider the slopes

$$\alpha_k := \frac{\widehat{T}(r_k,\varphi) - \widehat{T}(r_{k-1},\varphi)}{\log r_k - \log r_{k-1}}. \tag{7.1.7}$$

Assume that $\alpha_l \in [0,m]$. Then we have the following improvement over the bound (7.0.3):

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) - \frac{1}{m}\sum_{k=1}^{l} \frac{\left(\widehat{T}(r_k,\varphi)-\widehat{T}(r_{k-1},\varphi)\right)^2}{\log r_k - \log r_{k-1}}}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \tag{7.1.8}$$

By Lemma 7.1.4, the bound of Theorem 7.1.6 is only improved if one refines the sequence of subradii, subject to the inequality $\alpha_l \leq m$ on the slopes. Thus, in the limit, we obtain a continuous version of this theorem (see Theorem 7.1.10 below). In our experience, the extra numerical saving obtained in the limit is negligible once one chooses just a few division points. Moreover, it seems more practical from a computational standpoint to compute bounds on specific values of $\widehat{T}(r,\varphi)$ rather than integrals in terms of this function.

In order to formulate the continuous version of Theorem 7.1.6, we firstly introduce the following positive increasing function[28] of $r \in (0,1]$:

$$\widehat{A}(r,\varphi) := r\frac{d}{dr}\widehat{T}(r,\varphi), \tag{7.1.9}$$

whose notation mirrors the traditional covering spherical area function

$$\mathring{A}(r,\varphi) := \frac{1}{\pi}\iint_{D(0,r)} \frac{|\varphi'|^2}{(1+|\varphi|)^2}\,dxdy = \iint_{D(0,r)} \varphi^*\omega_{\mathrm{FS}} =: r\frac{d}{dr}\mathring{T}(r,\varphi)$$

of the Ahlfors–Shimizu theory.

[28]This is, in fact, a continuous function. We will not use this, and we do not give a proof of the C^1 property of $\widehat{T}(r,\varphi)$. For the abstract purpose (not used elsewhere in our paper, neither) of making an *almost everywhere* sense of the d/dr derivative in (7.1.9) and the $r \in [0,1]$ Riemann integral in (7.1.11), it suffices to appeal to Lebesgue's theorem that a monotone function $[0,1] \to \mathbf{R}$ is differentiable almost everywhere.

Now by interpreting the numerator of equation (7.1.8) as a Riemann sum, in the limit (under the assumption $\widehat{A}(r,\varphi) \leq \widehat{A}(1,\varphi) \leq m$ for $r \leq 1$) we obtain the following:

Theorem 7.1.10. *Assume the same conditions and notation as in Theorem 7.0.1. Assume that $\widehat{A}(1,\varphi) \leq m$. Then*

$$\begin{aligned} m &\leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) - \frac{1}{m}\int_0^1 \widehat{A}(r,\varphi)^2\,\frac{dr}{r}}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} \\ &= \frac{\int_{\mathbf{T}} \log|\varphi|\,\mu_{\mathrm{Haar}} + \int_0^1 \widehat{A}(r,\varphi)\,\frac{dr}{r} - \frac{1}{m}\int_0^1 \widehat{A}(r,\varphi)^2\,\frac{dr}{r}}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \end{aligned} \tag{7.1.11}$$

A further improvement is given in Theorem 7.6.4 by using a (heuristically speaking) optimal choice of the Euclidean metric in the evaluation module E_D.

7.1.12. *Binomial metrics.* As in Theorem 6.0.2, the set of analytic maps φ_n used to estimate the n^{th} archimedean evaluation height does not need to be of the particular form $\varphi_n(z) = \varphi(r(n)z)$. We include here one more elementary and fully explicit bound using $l+1 = 2$ knots with $\varphi_0(z) = Rz$ and $\varphi_1(z) = \varphi(z)$, but — unlike with § 7.1.1 — taking a family of metrics independent of the map φ. On the space of real auxiliary polynomials $E_D \otimes_{\mathbf{Z}} \mathbf{R} = \mathbf{R}[x]_{<D}$, the metric can be described as diagonalizing the monomials basis $\{x^k\}_{k=0}^{D-1}$ and giving the weights $\|x^k\| := \exp\left(\lambda D (k/D)^r + \mu k\right)$. The bound works out to the following explicit form, in which the triple of *binomial metric weights* $\{r,\lambda,\mu\}$ is to be optimized. Unlike for all our other holonomy bounds in this section § 7 and the next § 8, the proof of this theorem does not require any of the results from [BC22].

Theorem 7.1.13. *Assumptions and notation as in Theorem 7.0.1. We further assume that $f_1,\dots,f_m \in \mathbf{Q}[\![x]\!]$ are all convergent on the complex disc $|x| < R$. For $r \in \mathbf{R}_{>1}, \lambda \in \mathbf{R}_{>0}, \mu \in \mathbf{R}$, set*

$$\Gamma(x;r,\lambda,\mu) := \min\left\{ (r-1)\frac{\left(\max\{0, x-\mu\}\right)^{r/(r-1)}}{(r^r\lambda)^{1/(r-1)}}, \max\{(r-1)\lambda, x-\lambda-\mu\} \right\},$$

$$T(\varphi;r,\lambda,\mu) := \int_{\mathbf{T}} \Gamma(\log|\varphi(z)|;r,\lambda,\mu)\,\mu_{\mathrm{Haar}}(z),$$

$$T_{r,\lambda,\mu}(\varphi) := \frac{\lambda}{r+1} + \frac{\mu}{2} + T(\varphi;r,\lambda,\mu),$$

$$\chi_0 := \min\left\{1, \left(\frac{\max\{0, \log R - \mu\}}{\lambda r}\right)^{1/(r-1)}\right\}.$$

Suppose that $\mu \leq \log R < \log|\varphi'(0)|$ and

$$\chi_0 \leq \chi_1 := \frac{T(\varphi;r,\lambda,\mu) - \Gamma(\log R;r,\lambda,\mu)}{\log|\varphi'(0)| - \log R} \leq m, \qquad \chi_0 < 1.$$

Then

$$m \leq \frac{2T_{r,\lambda,\mu}(\varphi) - \frac{2}{m}\left(\frac{1}{2}\chi_1^2 \log\frac{|\varphi'(0)|}{R} + \chi_0\Gamma(\log R;r,\lambda,\mu) - \chi_0^2(\log R-\mu)\left(\frac{1}{2} - \frac{1}{r(r+1)}\right)\right)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \tag{7.1.14}$$

The extra assumption in this theorem concerning the inequalities among $\log R$, μ, $\log|\varphi'(0)|$, m, χ_1, χ_0 can be bypassed in the proof that follows to get, on a case-by-case basis, *some* bound on m in every case. We pick these particular conditions as they are satisfied in our applications. See Example 7.5.9.

The function $\Gamma(x;r,\lambda,\mu)$ emerges as the *Legendre transform* [Ell06, § VI] of the binomial metric weight function $\lambda t^r + \mu t$. This basic explicit computation is the content of our next lemma.

Lemma 7.1.15. *For arbitrary $r \in \mathbf{R}_{>1}, \lambda \in \mathbf{R}_{>0}, \mu \in \mathbf{R}$, and $x \in \mathbf{R}$, we have*

$$\max_{0\le t\le 1}\{tx - \lambda t^r - \mu t\} = \Gamma(x;r,\lambda,\mu).$$

Therefore,

$$T(\varphi;r,\lambda,\mu) = \int_{\mathbf{T}} \max_{0\le t\le 1}\{t\log|\varphi(z)| - \lambda t^r - \mu t\}\ \mu_{\mathrm{Haar}}(z).$$

Proof. The proof is a direct computation. Set $F(t) := xt - \lambda t^r - \mu t$, regarding x, λ, and μ as fixed. Then $F'(t) = x - r\lambda t^{r-1} - \mu$, and so F admits a critical point in $t \in \mathbf{R}_{\ge 0}$ if and only if $x - \mu \ge 0$, in which case the unique such critical point is

$$t_0 := \left(\frac{x-\mu}{r\lambda}\right)^{1/(r-1)}.$$

Therefore

$$\max_{0\le t\le 1} F(t) = \begin{cases} F(0) = 0 & \text{if } x-\mu \le 0 \\ F(t_0) = (r-1)\frac{(x-\mu)^{r/(r-1)}}{(r^r\lambda)^{1/(r-1)}} & \text{if } 0 \le x-\mu \le \lambda r \\ F(1) = x - \lambda - \mu & \text{if } x - \mu \ge \lambda r \end{cases}.$$

This is why we defined $\Gamma(x;r,\lambda,\mu)$ the way we did in the statement of Theorem 7.1.13. ☐

7.2. **A brief review of Bost's slopes method.** We review Bost's slopes method and related background material from Arakelov theory. The main references are [Bos01, §§4.1, 4.2] and [Bos20, Chapter 1]. For simplicity, we only recall the theory over $\mathbf{Q}$ as that is sufficient for our applications. Everything in this section holds verbatim for any number field, see [CDT24] and Remark 8.2.42.

7.2.1. *Hermitian vector bundles on* $\operatorname{Spec}\mathbf{Z}$.

Definition 7.2.2. A *Euclidean lattice* is a pair $\overline{E} = (E, \|\cdot\|)$ made of a finite rank free $\mathbf{Z}$-module E and a Euclidean norm $\|\cdot\|$ on the vector space $E_{\mathbf{R}}$. In other words: $\|\cdot\|^2$ is a positive definite quadratic form on $E_{\mathbf{R}} := E \otimes_{\mathbf{Z}} \mathbf{R}$.

In Arakelov geometry, this coincides with the notion of a Hermitian vector bundle on $\operatorname{Spec}\mathbf{Z}$. We thus use the notion of the *arithmetic degree* defined as the negative of the logarithm of the covolume of the Euclidean lattice:

$$\widehat{\deg}\,\overline{E} := -\log \operatorname{covol}(E, \|\cdot\|) = -\frac{1}{2}\log\left|\det\left(\langle e_i, e_j\rangle\right)_{i,j=1}^r\right|. \tag{7.2.3}$$

Here, $r := \operatorname{rank} E$,

$$\langle e, f\rangle := \frac{1}{2}\|e+f\|^2 - \|e\|^2 - \|f\|^2$$

is the associated inner product giving the quadratic form $\|e\| = \sqrt{\langle e,e\rangle}$, and $e_1,\dots,e_r$ is any $\mathbf{Z}$-module basis of E.

Let $M_{\mathbf{Q}}$ denote the equivalence classes (*places*) of absolute values $\mathbf{Q}\to[0,\infty)$. At the place $v\in M_{\mathbf{Q}}$, we select the representative absolute value $|\cdot|_v$ with the usual normalizations: $|\cdot|_\infty$ is the usual absolute value for the archimedean place $\infty\in M_{\mathbf{Q}}$, and $|p|_p = 1/p$ for the p-adic place $p\in M_{\mathbf{Q}}$. Thus $\prod_{v\in M_{\mathbf{Q}}}|x|_v = 1$ for all $x\in\mathbf{Q}^\times$. Let $M_{\mathbf{Q}}^{\mathrm{fin}} := M_{\mathbf{Q}}\smallsetminus\{\infty\}$ denote the set of all finite places of $\mathbf{Q}$, which is identified with the set of rational primes.

Along with the quadratic form $\|\cdot\|$ on $E_{\mathbf{R}}$, it is convenient to consider the p-adic norms $\|\cdot\|_p$ defined on $E_{\mathbf{Q}_p}$ by

$$\Big\|\sum_{i=1}^r c_ie_i\Big\|_p := \max_{1\le i\le r}|c_i|_p.$$

Note that $\|\cdot\|_p$ is independent of the choice of the basis $e_1,\dots,e_r$ of E: more intrinsically, we have $p^{\mathbf{Z}}$ for the value group of $\|\cdot\|_p$, with $\|w\|_p = p^{-n}$ if and only if $w\in E\otimes_{\mathbf{Z}}p^n\mathbf{Z}_p$ and $w\notin E\otimes_{\mathbf{Z}}p^{n+1}\mathbf{Z}_p$. Thus the Euclidean structure combines with the integral lattice structure E of the $\mathbf{Q}$-vector space $E_{\mathbf{Q}} := E\otimes_{\mathbf{Z}}\mathbf{Q}$ to define an adelic metric $(E_{\mathbf{Q}},(\|\cdot\|_v)_{v\in M_{\mathbf{Q}}})$. Conversely, we can recover the lattice $E\subset E_{\mathbf{Q}}\subset E_{\mathbf{R}}$ as the $w\in E_{\mathbf{Q}}$ defined by the simultaneous conditions $\|w\|_p\le 1$ for all primes p.

In these notations, the arithmetic degree formula (7.2.3) takes the following adelic form:

$$\widehat{\deg}\,\overline{E} = -\frac{1}{2}\log\left|\det\left(\langle v_i,v_j\rangle\right)_{i,j=1}^r\right| - \sum_{p\in M_{\mathbf{Q}}^{\mathrm{fin}}}\sum_{i=1}^r\log\|v_i\|_p,$$

where $\{v_1,\dots,v_r\}$ is any $\mathbf{Q}$-basis of $E_{\mathbf{Q}}$.

Given two Euclidean lattices $\overline{E},\overline{F}$, let $\overline{E}\oplus\overline{F}$ denote $E\oplus F$ equipped with the norm given by the orthogonal direct sum of the norms on the subspaces $E_{\mathbf{R}}$ and $F_{\mathbf{R}}$. By definition, we have

$$\widehat{\deg}\,(\overline{E}\oplus\overline{F}) = \widehat{\deg}\,(\overline{E}) + \widehat{\deg}\,(\overline{F}). \tag{7.2.4}$$

7.2.5. *Slopes of Euclidean lattices and heights of morphisms.*

Definition 7.2.6. The *slope* $\widehat{\mu}(\overline{E})$ of a Euclidean lattice $\overline{E} = (E,\|\cdot\|)$ is defined as

$$\widehat{\mu}(\overline{E}) := \frac{\widehat{\deg}\,\overline{E}}{\operatorname{rank}E}\in\mathbf{R}.$$

The *maximal slope* of $\overline{E}$ is defined as

$$\widehat{\mu}_{\max}(\overline{E}) := \sup_{0\subsetneq F\subseteq E}\widehat{\mu}(\overline{F}),$$

where F runs through all nonzero $\mathbf{Z}$-submodules of E and $\overline{F}$ denotes the induced Euclidean lattice of F equipped with the quadratic form obtained from restricting $\|\cdot\|$ to $F_{\mathbf{R}}$.

The following lemma follows from the definition.

Lemma 7.2.7. *Let $\overline{E}, \overline{F}$ be two Euclidean lattices. Let $\overline{E} \otimes \overline{F}$ denote $E \otimes_{\mathbf{Z}} F$ equipped with the tensor norm. Then we have*

$$\widehat{\mu}(\overline{E} \otimes \overline{F}) = \widehat{\mu}(\overline{E}) + \widehat{\mu}(\overline{F}).$$

Proof. See, for instance, [Che09, Lemma 2.3], which we briefly summarize here. By definition of the arithmetic degree, we have $\widehat{\deg}(\overline{E}) = \widehat{\deg}(\wedge^{\operatorname{rank} E}\overline{E})$, and for any two rank 1 Euclidean lattices $\overline{L}_1, \overline{L}_2$, we have $\widehat{\deg}(\overline{L}_1 \otimes \overline{L}_2) = \widehat{\deg}(\overline{L}_1) + \widehat{\deg}(\overline{L}_2)$. Note that

$$\wedge^{\operatorname{rank} E \otimes F}(\overline{E} \otimes \overline{F}) \cong (\wedge^{\operatorname{rank} E}\overline{E})^{\otimes \operatorname{rank} F} \otimes (\wedge^{\operatorname{rank} F}\overline{F})^{\otimes \operatorname{rank} E}.$$

Thus

$$\widehat{\deg}(\overline{E} \otimes \overline{F}) = (\operatorname{rank} F)\,\widehat{\deg}(\overline{E}) + (\operatorname{rank} E)\,\widehat{\deg}(\overline{F}),$$

completing the proof of the lemma. □

Consider two Euclidean lattices $\overline{E} = (E, \|\cdot\|_E)$ and $\overline{F} = (F, \|\cdot\|_F)$ and an *injective* homomorphism $\psi : E_{\mathbf{Q}} \hookrightarrow F_{\mathbf{Q}}$. If ψ sends the Euclidean lattice E isometrically into a sublattice of F, then $\widehat{\mu}(\overline{E}) \leq \widehat{\mu}_{\max}(\overline{F})$ by the definition of the maximal slope. In general, the slope of the source lattice $\overline{E}$ can be upper-estimated in terms of the maximal slope of the range lattice $\overline{F}$ and the *height* of the homomorphism ψ.

Definition 7.2.8. The *local v-adic height* (at a place $v \in M_{\mathbf{Q}}$) of the monomorphism ψ is defined as the logarithm of the norm of the induced monomorphism

$$\big(E_{\mathbf{Q}_v}, \|\cdot\|_{E,v}\big) \hookrightarrow \big(F_{\mathbf{Q}_v}, \|\cdot\|_{F,v}\big)$$

of normed $\mathbf{Q}_v$-vector spaces:

$$h_v(\psi) := \sup_{e \in E_{\mathbf{Q}_v} \smallsetminus \{0\}} \log \frac{\|\psi(e)\|_{F,v}}{\|e\|_{E,v}} = \sup_{e \in E \smallsetminus \{0\}} \log \frac{\|\psi(e)\|_{F,v}}{\|e\|_{E,v}}.$$

The *global height* of ψ is the sum of the local v-adic heights over all places $v \in M_{\mathbf{Q}}$:

$$h(\psi) := \sum_{v \in M_{\mathbf{Q}}} h_v(\psi).$$

The tautological inequality $\widehat{\mu}(\overline{E}) \leq \widehat{\mu}_{\max}(\overline{F})$ for the isometric injections $E \hookrightarrow F$ then generalizes to arbitrary monomorphisms $\psi : E_{\mathbf{Q}} \hookrightarrow F_{\mathbf{Q}}$, in the following way.

Lemma 7.2.9 ([Bos01], Prop. 4.5)**.** *For every monomorphism*

$$\psi : E_{\mathbf{Q}} \hookrightarrow F_{\mathbf{Q}}$$

of the induced $\mathbf{Q}$-vector spaces of the Euclidean lattices $\overline{E}$ and $\overline{F}$, we have

$$\widehat{\mu}(\overline{E}) \leq \widehat{\mu}_{\max}(\overline{F}) + h(\psi). \tag{7.2.10}$$

7.2.11. *Bost's slopes inequality.* For filtered Euclidean lattices, the slopes inequality (7.2.10) generalizes as follows. Let F be a free $\mathbf{Z}$-module, which we no longer require to be of finite rank. We suppose that there is a filtration on $F_{\mathbf{Q}}$

$$F^{\bullet}_{\mathbf{Q}} \,:\, F_{\mathbf{Q}} = F^{(0)}_{\mathbf{Q}} \supseteq F^{(1)}_{\mathbf{Q}} \supseteq F^{(2)}_{\mathbf{Q}} \supseteq \cdots$$

with finite-dimensional graded quotients $\operatorname{Gr}_n(F^{\bullet}_{\mathbf{Q}}) := F^{(n)}_{\mathbf{Q}}/F^{(n+1)}_{\mathbf{Q}}$ and such that $\cap_{n=0}^{\infty} F^{(i)}_{\mathbf{Q}} = \{0\}$.

We consider $\overline{E}, \psi$ as in § 7.2.5. In particular, the linear monomorphism $\psi : E_{\mathbf{Q}} \hookrightarrow F_{\mathbf{Q}}$ induces a filtration on E:

$$E^\bullet \,:\, E = E^{(0)} \supseteq E^{(1)} \supseteq \cdots, \quad \text{where } E^{(n)} = E \cap \psi^{-1}(F_{\mathbf{Q}}^{(n)}).$$

Note that since E is a finite rank free $\mathbf{Z}$-module and ψ is injective, we have that $\mathrm{Gr}_n(E^\bullet)$ are finite rank free $\mathbf{Z}$-modules and the above filtration stabilizes to $\{0\}$ after finitely many steps. Moreover, the restriction of $\|\cdot\|_E$ to $E^{(n)}$ gives $E^{(n)}$ a Euclidean lattice structure and the corresponding quotient metric on $E^{(n)}/E^{(n+1)}$ equipped it with a Euclidean lattice structure $\overline{E^{(n)}/E^{(n+1)}}$.

We also assume that each graded quotient piece $\mathrm{Gr}_n(F_{\mathbf{Q}}^\bullet)$ is endowed with a Euclidean lattice structure. More precisely, for each n, we have a Euclidean lattice

$$\overline{G^{(n)}} = \big(G^{(n)}, \|\cdot\|_{G^{(n)}}\big),$$

where $G^{(n)} \subset \mathrm{Gr}_n(F_{\mathbf{Q}}^{(n)})$ a $\mathbf{Z}$-submodule such that $G_{\mathbf{Q}}^{(n)} = \mathrm{Gr}_n(F_{\mathbf{Q}}^{(n)})$.

The map ψ induces a linear monomorphism between the graded quotients:

$$\psi_D^{(n)} \,:\, \mathrm{Gr}_n(E^\bullet)_{\mathbf{Q}} \hookrightarrow \mathrm{Gr}_n(F_{\mathbf{Q}}^\bullet) \tag{7.2.12}$$

and its height $h(\psi_D^{(n)})$ is defined using the above-mentioned Euclidean lattices structures $\overline{E^{(n)}/E^{(n+1)}}$ and $\overline{G^{(n)}}$.

Lemma 7.2.13 ([Bos01], Prop. 4.6)**.** *In this situation,*

$$\widehat{\deg}\,(\overline{E}) \leq \sum_{n=0}^{\infty} \mathrm{rank}\,(E^{(n)}/E^{(n+1)}) \Big[\widehat{\mu}_{\max}(\overline{G^{(n)}}) + h(\psi_D^{(n)})\Big]. \tag{7.2.14}$$

Note that the above sum is a finite sum since $E^{(N)} = 0$ for $N \gg 1$.

7.3. **The Bost–Charles bound.** We follow [Bos20, BC22] with a slight modification to take denominators into account.

We recall the setting of their work for our application. Consider $\mathcal{X} = \mathbf{P}^1_{\mathbf{Z}}$ and the line bundle $\mathcal{L} := \mathcal{O}(1)$ on $\mathcal{X}$. We denote by $x := X_1/X_0$ the coordinate of (an affine line in) $\mathbf{P}^1_{\mathbf{Z}} = \mathrm{Proj}\,\mathbf{Z}[X_0, X_1]$, and then for $D \in \mathbf{Z}_{>0}$ we follow the usual identifications $\mathcal{L} = \mathcal{O}([0])$ (here $[0]$ denotes the divisor of the point $x = 0$) and $\mathbf{Z}[1/x]_{\leq D} = \Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})$.

7.3.1. *The Bost–Charles metric.* Following the ideas in [BC22, §§8.2, 8.3], using $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{P}^1(\mathbf{C}), 0)$, we endow $\mathcal{L} = \mathcal{O}(1)$ with the Hermitian metric $\|\cdot\|_{\overline{\mathcal{L}}}$ defined by

$$\|\mathbf{1}(y)\|_{\overline{\mathcal{L}}} := \exp\left(-\sum_{z \in \varphi^{-1}(y)} \log^+ \frac{1}{|z|}\right) = \prod_{z \in \overline{\mathbf{D}},\, \varphi(z) = y} |z|,$$

where $\mathbf{1} = \mathbf{1}_{[0]}$ is the canonical section ("constant function") of $\mathcal{L} = \mathcal{O}([0])$ corresponding to the divisor $[0]$, and $y \in \mathbf{P}^1(\mathbf{C})$. This Hermitian metric has $\mathcal{C}^{\mathrm{b}\Delta}$ regularity in the sense of Bost–Charles (see [BC22, §§4.1.1, 4.2.1.2]). We shall denote this Hermitian line bundle by $\overline{\mathcal{O}(1)}$, or by $\overline{\mathcal{O}(1)}_\varphi$ if we wish to indicate the dependence on φ. We work in the framework of arithmetic intersection theory using such Hermitian line bundles and Arakelov divisors with $\mathcal{C}^{\mathrm{b}\Delta}$ regularity in the sense of Bost–Charles [BC22, §4.5].

Remark 7.3.2. In [BC22], the Hermitian metric of § 7.3.1 is given by the Arakelov divisor $([0], \varphi_*\left(\log^+|z|^{-1}\right))$. More precisely, in the sense of [BC22, §6.2.1 using Example 4.3.1], we have the compactly supported Arakelov divisor $([0], \log^+|z|^{-1})$ on the smooth formal-analytic (hereafter, f.-a.) arithmetic surface

$$\widetilde{\mathcal{V}}(\varphi) := (\operatorname{Spf} \mathbf{Z}[\![x]\!], (\overline{\mathbf{D}}, 0), i_\varphi)$$

over $\mathbf{Z}$, where $x \mapsto \varphi$ defines an isomorphism $\mathbf{C}[\![x]\!] \xrightarrow{\simeq} \mathbf{C}[\![z]\!]$ (here z denotes the coordinate on $\overline{\mathbf{D}}$, and we use the assumption that $\varphi'(0) \neq 0$), and thus its compositional inverse induces an isomorphism $i_\varphi : \operatorname{Spf} \mathbf{C}[\![x]\!] \cong \widehat{\mathbf{D}}_0$. See [Bos20, §10.6.1] or [BC22, §6.1.1] for the general definition of smooth f.-a. arithmetic surface over a number field, and § 6.4.1.1 in *loc. cit.* for this construction $\widetilde{\mathcal{V}}(\varphi)$, which also comes with a distinguished nonconstant regular function $(\iota, \varphi) : \widetilde{\mathcal{V}}(\varphi) \to \mathbf{A}^1_{\mathbf{Z}}$ on the f.-a. arithmetic surface; cf [BC22, §7.1.1.1]. Here, $\iota : \operatorname{Spf} \mathbf{Z}[\![x]\!] \hookrightarrow \operatorname{Spec} \mathbf{Z}[x] = \mathbf{A}^1_{\mathbf{Z}}$ is the natural formal immersion. In the setting of [BC22, §7.2.1], the Arakelov divisor $([0], \varphi_*(\log^+|z|^{-1}))$ is the direct image of $([0], \log^+|z|^{-1})$ by the morphism $(\iota, \varphi) : \widetilde{\mathcal{V}}(\varphi) \to \mathbf{A}^1_{\mathbf{Z}}$, where the pushforward map φ_* on Green functions is defined in [BC22, §3.4.2.1]. By [BC22, Corollary 4.4.2(ii)], this pushforward map preserves $\mathcal{C}^{b\Delta}$ regularity of Green functions, which is essential for having a well-behaved arithmetic intersection theory. The explicit formula of the Hermitian metric associated to $\varphi_*(\log^+|z|^{-1})$ is given in [BC22, §5.1.2]. △

7.3.3. *Direct images and arithmetic Hilbert–Samuel.* Fix a smooth probability measure ν on $\mathbf{P}^1(\mathbf{C})$, for instance the Fubini–Study form $\omega_{\mathrm{FS}} = \frac{\sqrt{-1}}{2\pi}\frac{dz \wedge d\bar{z}}{(1+|z|^2)^2}$; the choice of ν is immaterial to the proof. As in [BC22, § 6], the Hermitian metric on $\overline{\mathcal{L}}$ combines with fiberwise integration over the manifold $\mathcal{X}(\mathbf{C})$ to define a Euclidean lattice structure on the $\mathbf{Z}$-module $\Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})$. Explicitly, we norm $s \in \Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})$ by

$$\|s\| := \sqrt{\int_{\mathcal{X}(\mathbf{C})} \|s\|^2_{\overline{\mathcal{L}}}\, \nu}.$$

Following [BC22, § 6.1.2.2], we denote by $\Gamma_{L^2}\left(\mathcal{X}, \nu; \overline{\mathcal{L}}^{\otimes D}\right)$ this Euclidean lattice. Up to the integration metric weight ν, in a $D \to \infty$ asymptotic sense, this is essentially the zeroth direct image of $\overline{\mathcal{L}} = \overline{\mathcal{O}}(1)_\varphi$ under the structure morphism $\widetilde{\mathcal{V}}(\varphi) \to \operatorname{Spec} \mathbf{Z}$. As in [BC22, §8, Theorem 8.2.5], we can express the arithmetic Hilbert–Samuel formula on the arithmetic surface $\mathcal{X} = \mathbf{P}^1_{\mathbf{Z}}$ into the form

$$\widehat{\deg}\, \Gamma_{L^2}\left(\mathcal{X}, \nu; \overline{\mathcal{L}}^{\otimes D}\right) = \frac{1}{2}(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}})D^2 + o(D^2). \tag{7.3.4}$$

When the Hermitian metric in $\mathcal{L} = \mathcal{O}(1)$ is smooth, this formula is due to Zhang [Zha95, Theorem 1.4] with an additional input by Bost in comparing two Hermitian metrics in the proof of [Bos20, Theorem 10.3.2]. Zhang's theorem is a refinement to (non-pointwise-strict) semipositive curvature (Chern form) $c_1(\mathcal{L}, \|\cdot\|) \geq 0$ of the work of Gillet–Soulé [GS92] and Bismut–Vasserot [BV89]; see also Abbes–Bouche [AB95] for an outline of a more direct approach. Following the idea in [Bos99, §5] and [BC22, §§3–4] to separate the Green function into a smooth Green function and a $\mathcal{C}^{b\Delta}$ function, the same arithmetic Hilbert–Samuel formula holds for ample line bundles with $\mathcal{C}^{b\Delta}$ Hermitian metrics of pointwise non-negative Chern form (as defined in [BC22, § 4.2.1.2]), and so the formula is also valid in our setting.

In terms of φ, Bost and Charles [BC22, Theorem 5.4.1 and Proposition 5.4.2] provide the following formula for the self-intersection number:

$$(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}) = \left(\overline{\mathcal{O}(1)}_{\varphi}\cdot\overline{\mathcal{O}(1)}_{\varphi}\right) = \iint_{\mathbf{T}^2}\log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\,\mu_{\mathrm{Haar}}(w) = \widehat{T}(1,\varphi). \tag{7.3.5}$$

We will review their computation in our mild generalization in Lemma 7.4.5 further down.

Proof of Theorem 7.0.1. Note that the choice of $\mathbf{b}$ is not unique; we may permute the columns of $\mathbf{b}$ without changing the form of the f_i. Therefore, after a suitable permutation of the columns, we may assume $u_1 \le u_2 \le \ldots \le u_r$. *We keep this convention for all the proofs in § 7.*

For $D \in \mathbf{N}$, we take for our evaluation module the following free $\mathbf{Z}$-module of rank $m(D+1)$:

$$E_D := \bigoplus_{h=0}^{r}\bigoplus_{i=u_h+1}^{u_{h+1}} \frac{[1,\ldots,\xi D]^{e_i}}{[1,\ldots,y_{h+1}D]\cdots[1,\ldots,y_rD]}\, f_i\cdot\mathbf{Z}[1/x]_{\le D}, \tag{7.3.6}$$

where $u_0 := 0, u_{r+1} := m$, $\xi \in [0,m]$ and $y_h \in [0, b_h m] \subset \mathbf{R}$ are auxiliary parameters to be optimized in the proof. Note that the indexing of E_D in equation (7.3.6) differs slightly from the notation of § 2.13.1; the difference amounts to considering polynomials of degree $< D$ rather than $\le D$. We use this normalization — which asymptotically makes no difference and so is ultimately an aesthetic choice — so that we can talk below about sections of $\mathcal{L}^{\otimes D}$ rather than $\mathcal{L}^{\otimes(D-1)}$.

To endow E_D with a Euclidean norm, we take the orthogonal direct sum (7.3.6) of the lattices

$$\frac{[1,\ldots,\xi D]^{e_i}}{[1,\ldots,y_{h+1}D]\cdots[1,\ldots,y_rD]}\mathbf{Z}[1/x]_{\le D} \subset \Gamma\left(\mathcal{X},\mathcal{L}^{\otimes D}\right)_{\mathbf{R}},$$

with each of these summands inheriting the norm induced from $\overline{\mathcal{O}(1)}_{\varphi} = \overline{\mathcal{L}}$. We shall denote this Euclidean lattice by $\overline{E}_D$.

By (7.3.4) and (7.3.5), we have

$$\begin{aligned}\widehat{\deg}\,\overline{E}_D &= \left(\frac{m}{2}\left(\overline{\mathcal{O}(1)}_{\varphi}.\overline{\mathcal{O}(1)}_{\varphi}\right) + \sum_{h=1}^{r} u_h y_h - \xi\left(\sum_{i=1}^{m} e_i\right)\right)D^2 + o(D^2)\\ &= \left(\frac{m}{2}\widehat{T}(1,\varphi) + \sum_{h=1}^{r} u_h y_h - \xi\Big(\sum_{i=1}^{m} e_i\Big)\right)D^2 + o(D^2).\end{aligned} \tag{7.3.7}$$

Let X denote $\mathcal{X}_{\mathbf{Q}} = \mathbf{P}^1_{\mathbf{Q}}$. We identify $\operatorname{Spf}\mathbf{Q}[\![x]\!] = \widehat{X}_0$ as the formal completion of X at its closed subscheme 0. This designates $f_i(x) \in \Gamma(\widehat{X}_0, \mathcal{O}_{\widehat{X}_0})$, for $i = 1,\ldots,m$. Let $\Gamma(\widehat{X}_0,\mathcal{L}^{\otimes D})$ denote the global sections of $\mathcal{L}^{\otimes D}|_{\widehat{X}_0}$; these get identified with

$$\Gamma(\widehat{X}_0,\mathcal{L}^{\otimes D}) = x^{-D}\mathbf{Q}[\![\mathbf{x}]\!] =: F_{\mathbf{Q}}.$$

(Here, $x^{-D}\mathbf{Q}[\![\mathbf{x}]\!]$ denotes the $\mathbf{Q}$-vector space generated by x^k, where $k \ge -D$.) Thus $f_i\Gamma(\mathcal{X},\mathcal{L}^{\otimes D}) \subset \Gamma(\widehat{X}_0,\mathcal{L}^{\otimes D})$, and we have the evaluation map

$$\psi_D : E_D\otimes_{\mathbf{Z}}\mathbf{Q} \hookrightarrow F_{\mathbf{Q}}, \qquad (Q_i)_{1\le i\le m} \mapsto \sum_{i=1}^{m} f_i Q_i, \tag{7.3.8}$$

where $Q_i \in \Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})_{\mathbf{Q}}$. It is an injective homomorphism due to our assumed $\mathbf{Q}(x)$-linear independence of the formal functions $f_i(x)$.

We filter F by the $x = 0$ vanishing order of the formal sections of $\mathcal{L}^{\otimes D}|_{\widehat{X}_0}$:

$$F_{\mathbf{Q}} = F_{\mathbf{Q}}^{(0)} \supseteq F_{\mathbf{Q}}^{(1)} \supseteq \cdots \supseteq F_{\mathbf{Q}}^{(n)} \supseteq \cdots .$$

Concretely, $F_{\mathbf{Q}}^{(n)} = \mathrm{Span}_{\mathbf{Q}}\{x^{k-D} \mid k \geq n\}$. The graded piece $F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}$ is a one dimensional $\mathbf{Q}$-vector space generated by the image of x^{n-D} under the quotient map. We take the Euclidean lattice structure on $F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}$ given by the free rank one $\mathbf{Z}$-module generated by the image of x^{n-D} and the Euclidean norm with $\|x^{n-D}\| = 1$. Note that these Euclidean lattice structures on graded piece are all induced from the free $\mathbf{Z}$-module $F = x^{-D}\mathbf{Z}[\![x]\!]$ and the Euclidean norm on $x^{-D}\mathbf{R}[x]$ that has $\{x^n\}_{n \in \mathbf{Z}_{\geq -D}}$ for an orthonormal basis.

As in § 7.2.11, we use $E_D^{(n)} := \psi_D^{-1}\left(F_{\mathbf{Q}}^{(n)}\right) \cap E_D$ to denote the preimage of $F_{\mathbf{Q}}^{(n)}$ in E_D under ψ_D. For each $n \in \mathbf{N}$, the evaluation map (7.3.8) induces an injective homomorphism

$$\psi_D^{(n)} : E_D^{(n)}/E_D^{(n+1)} \hookrightarrow F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}.$$

In particular, as in (3.1.4), we have $\mathrm{rank}\left(E_D^{(n)}/E_D^{(n+1)}\right) \in \{0, 1\}$. Let

$$\mathcal{V}_D := \left\{n \in \mathbf{N} \mid \mathrm{rank}\left(E_D^{(n)}/E_D^{(n+1)}\right) = 1\right\}.$$

We have $\#\mathcal{V}_D = \mathrm{rank}\, E_D = m(D+1)$.

We now provide upper bounds on the evaluation heights $h_\infty(\psi_D^{(n)})$ and $h_{\mathrm{fin}}(\psi_D^{(n)})$, where $h_{\mathrm{fin}}(\psi_D^{(\mathbf{n})}) := \sum_{v \in M_{\mathbf{Q}}, v \nmid \infty} h_v(\psi_D^{(\mathbf{n})})$.

The archimedean evaluation height bound stems from the work of Bost and Bost–Charles:

$$h_\infty(\psi_D^{(n)}) \leq -n \log|\varphi'(0)| + \left(\overline{\mathcal{O}(1)}_\varphi \cdot \overline{\mathcal{O}(1)}_\varphi\right) D + o(D). \tag{7.3.9}$$

The proof details for our specific setting are in either § 7.4 (Lemma 7.4.1, specializing to $r = 1$) or in § 8.2.11) (specializing to $d = 1$), where respectively we will need a refinement of this estimate to incorporate convexity and handle the high dimensional setup. For the original source we refer to the two paragraphs following Theorem 8.2.2 on page 127 in [BC22], which in turn summarize the relevant sections of [Bos20]. The specific bound is essentially [Bos20, §10.5.5, Theorem 10.5.3, Corollary 10.5.4].

Next we estimate $h_{\mathrm{fin}}(\psi_D^{(n)})$. For each prime p, by the definition of h_p, our task is to consider an arbitrary element $(Q_i)_{1 \leq i \leq m} \in E_D^{(n)} \smallsetminus E_D^{(n+1)}$, and to provide an upper bound on $\log|c_n|_p$, where c_n denotes the (leading order n) coefficient of x^n in $\sum_{i=1}^m f_i Q_i = c_n x^n + \ldots$.

Recall the notation of the indices cutoffs $u_j \in \{0, 1, \ldots, m\}$, for $1 \leq j \leq r$, from the statement of Theorem 6.0.2. Let h_i be the index in $\{0, 1, \ldots, r\}$ defined by $u_{h_i} < i \leq u_{h_i+1}$. The ultrametric triangle inequality for $|\cdot|_p$ directly gives

$$\begin{aligned}
&\frac{\log |c_n|_p}{\log p} \\
&\le \max_{\substack{1\le i\le m\\ 0\le k\le \min\{n-1,D\}}} \left\{ \mathrm{val}_p\left(\prod_{j=1}^{h_i}[1,\dots,b_j(n-k)]\cdot\prod_{j=h_i+1}^{r}[1,\dots,y_jD]\right)+\mathrm{val}_p\left(\frac{(n-k)^{e_i}}{[1,\dots,\xi D]^{e_i}}\right)\right\} \\
&\le \left(\sum_{h=1}^{r}\mathrm{val}_p\left([1,\dots,b_h\max\{n,(y_h/b_h)D\}]\right)\right)+\left(\max_{1\le i\le m} e_i\right)\mathrm{val}_p\left(\frac{[\max\{n-D,1\},\dots,n]}{[1,\dots,\xi D]}\right).
\end{aligned} \tag{7.3.10}$$

Note that for $n^{1/2} < p \le \xi D$ we have $\mathrm{val}_p\left([\max\{n-D,1\},\dots,n]/[1,\dots,\xi D]\right)\le 0$. Since all terms under the p-adic valuation in (7.3.10) are independent of p, the prime number theorem gives

$$\begin{aligned}
h_{\mathrm{fin}}(\psi_D^{(n)}) &\le \sum_{h=1}^{r}\log([1,\dots,b_h\max\{n,(y_h/b_h)D\}]) \\
&\quad+\left(\max_{1\le i\le m} e_i\right)\sum_{p>\max\{n^{1/2},\xi D\}}\mathrm{val}_p([\max\{n-D,1\},\dots,n])\log p+o(n) \\
&\le \sum_{h=1}^{r} b_h\max\{n,(y_h/b_h)D\} \\
&\quad+\left(\max_{1\le i\le m} e_i\right)\sum_{\substack{p>\max\{n^{1/2},\xi D\},\\ p\mid[\max\{n-D,1\},\dots,n]}}\log p+o(n+D).
\end{aligned}$$

By Lemma 5.0.4, we have for $n\ge \max\{\xi,1\}D$

$$\begin{aligned}
h_{\mathrm{fin}}(\psi_D^{(n)}) \le \left(\max_{1\le i\le m} e_i\right)&\left(\left(D\sum_{j=1}^{\lfloor (n/D-1)/\max(1,\xi)\rfloor}1/j\right.\right. \\
&\left.\left.+\left(\frac{n}{\lfloor (n/D+(\xi-1)^+)/\max(1,\xi)\rfloor}-\xi D\right)^+\right)\right. \\
&+\sum_{h=1}^{r} b_h\max\{n,(y_h/b_h)D\}+o(n+D);
\end{aligned} \tag{7.3.11}$$

Again by Lemma 5.0.4 (taking $k=n-1$), we have for $\min\{\xi,1\}D\le n<D$

$$\begin{aligned}
h_{\mathrm{fin}}(\psi_D^{(n)}) \le \left(\max_{1\le i\le m} e_i\right)&(n-\xi D)^+ \\
&+\sum_{h=1}^{r} b_h\max\{n,(y_h/b_h)D\}+o(n+D);
\end{aligned} \tag{7.3.12}$$

for $n<\xi D$ the estimate is just

$$h_{\mathrm{fin}}(\psi_D^{(n)}) \le \sum_{h=1}^{r} b_h\max\{n,(y_h/b_h)D\}+o(n)+o(D). \tag{7.3.13}$$

André's Corollary 2.6.1, proved in Appendix B but also in the self-contained Lemma 7.3.17 below, permits us to apply the Chudnovsky–Osgood Theorem 3.2.13

on functional bad approximability. In the $D \to \infty$ asymptotic, by the continuity of $v \mapsto I_u^v(w)$, this entails the total evaluation height upper bounds

$$\sum_{n \in \mathcal{V}_D} h_\infty(\psi_D^{(n)}) \leq \left(-\frac{m^2}{2} \log |\varphi'(0)| + m(\overline{\mathcal{O}(1)} \cdot \overline{\mathcal{O}(1)}) \right) D^2 + o(D^2). \tag{7.3.14}$$

and

$$\begin{aligned}\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)}) &\leq \left(\left(\max_{1 \leq i \leq m} e_i \right) I_\xi^m(\xi) + \sum_{h=1}^r b_h \int_0^m \max\{s, y_h/b_h\}\, ds \right) D^2 + o(D^2) \\ &= \left(\left(\max_{1 \leq i \leq m} e_i \right) I_\xi^m(\xi) + \frac{1}{2} \left(\sigma_m m^2 + \sum_{h=1}^r y_h^2/b_h \right) \right) D^2 + o(D^2).\end{aligned} \tag{7.3.15}$$

Let us give more details on how to obtain (7.3.15) using Theorem 3.2.13; the verbatim reasoning applies also to (7.3.14) and to other similar evaluation height estimates in the rest of the section. Firstly, Lemma 7.3.17 and our standing assumptions in Theorem 7.0.1 imply that $f_1, \ldots, f_m$ are holonomic functions. Let ε and $C(\varepsilon)$ be as in the statement of Theorem 3.2.13. Throughout this section, the evaluation heights $h_\infty(\psi_D^{(n)})$ and $h_{\mathrm{fin}}(\psi_D^{(n)})$ get asymptotically upper-estimated with $o(D+n)$ implicit error terms, and with certain explicit nonnegative main terms that are, in all cases, certainly $\geq -C'D$, uniformly in $D \gg 1$ and ε; in these estimates, C' as well as the decay rates in the $o(D+n)$ of the error terms only depend on m, $\{f_i\}$, and φ. (For the $h_{\mathrm{fin}}(\psi_D^{(n)})$ situation under current highlight, we may of course simply take $C' = 0$; we keep C' to illustrate how the argument in the other situations.) For any $\varepsilon > 0$, Theorem 3.2.13 gives $\mathcal{V}_D \subset [0, (m+\varepsilon)(D+1)+C(\varepsilon)]$ with $\#\mathcal{V}_D = m(D+1)$. Thus the total evaluation height $\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)})$ is majorized by the $0 \leq n \leq (m+\varepsilon)(D+1)+C(\varepsilon)$ sum of (7.3.11), resp. (7.3.12), (7.3.13), minus the overcount of at most $\varepsilon(D+1)+C(\varepsilon)$ terms, to all of which we apply the $\geq -C'D$ lower bound to compensate. In the situation at hand, we get asymptotically

$$\begin{aligned}&\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)}) \\ &\leq \left(\left(\max_{1 \leq i \leq m} e_i \right) I_\xi^{m+\varepsilon+C(\varepsilon)/D}(\xi) + \sum_{h=1}^r b_h \int_0^{m+\varepsilon+C(\varepsilon)/D} \max\{s, y_h/b_h\}\, ds \right) D^2 \\ &+ C'D(\varepsilon(D+1) + C(\varepsilon)) + o(D^2).\end{aligned}$$

By the continuity of $I_u^v(w)$ in v, we derive with an arbitrary $\varepsilon > 0$ the upper estimate

$$\lim_{D \to \infty} \frac{\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)})}{D^2} \leq \left(\left(\max_{1 \leq i \leq m} e_i \right) I_\xi^{m+\varepsilon}(\xi) + \sum_{h=1}^r b_h \int_0^{m+\varepsilon} \max\{s, y_h/b_h\}\, ds \right) + C'\varepsilon.$$

As the left-hand side is independent of the choice of ε in Theorem 3.2.13, we can let $\varepsilon \to 0$ and obtain

$$\lim_{D \to \infty} \frac{\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)})}{D^2} \leq \left(\max_{1 \leq i \leq m} e_i \right) I_\xi^m(\xi) + \sum_{h=1}^r b_h \int_0^m \max\{s, y_h/b_h\}\, ds.$$

This is exactly (7.3.15).

Now Bost's slopes inequality (7.2.14) reads, in our situation:

$$\widehat{\deg}\,\overline{E}_D \leq \sum_{n \in \mathcal{V}_D} h_\infty(\psi_D^{(n)}) + \sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)}). \tag{7.3.16}$$

Combining the upper bounds (7.3.7), (7.3.14), and (7.3.15), and picking out the coefficients of the leading D^2 of the $D \to \infty$ asymptotic, we derive by Bost's inequality (7.2.14) the upper bound

$$(\log|\varphi'(0)| - \sigma_m)m^2 \leq m\left(\overline{\mathcal{O}(1)}_\varphi \cdot \overline{\mathcal{O}(1)}_\varphi\right) + 2\left(\xi\left(\sum_{i=1}^m e_i\right) + \left(\max_{1\leq i\leq m} e_i\right) I_\xi^m(\xi)\right) + \left(\sum_{h=1}^r y_h^2/b_h - 2\sum_{h=1}^r u_h y_h\right).$$

The quadratic form $\sum_{h=1}^r y_h^2/b_h - 2\sum_{h=1}^r u_h y_h$ reaches its minimum when $y_h = u_h b_h$ for all $1 \leq h \leq r$, and thus with (7.3.5) we obtained the desired bound. □

Lemma 7.3.17. *If $\log|\varphi'(0)| > \sigma_m$, then all the f_i are holonomic.*

(See also Corollary 2.6.1 and its proof in Appendix B.)

Proof. Applying the differential operator $\left(x\frac{d}{dx}\right)^{e_i}$ to remove the terms n^{e_i} from the denominators of the coefficients of f_i, we may and do assume — with no loss of generality for the goal of proving the present lemma — that $\mathbf{e} = \mathbf{0}$. The following then is the familiar calculation as in Appendix B, which in effect gives another proof of (B.0.1), now in the framework of Bost's slopes inequality. For our concrete purposes here, we need for every i to construct a $\mathbf{Q}(x)$-linear dependency among the derivatives $f_i, f_i', f_i'', \ldots$. Suppose to the contrary that all those derivatives are $\mathbf{Q}(x)$-linearly independent. With an arbitrary $m' \in \mathbf{N}$, we apply a stripped down form of the main argument of the present section, now to the rank-$(m'+1)D$ evaluation module

$$E_D = \bigoplus_{j=0}^{m'} f_i^{(j)} \mathbf{Z}[x]_{<D},$$

equipped with the Euclidean norm in which $\{f_i^{(j)} x^k\}$ is an orthonormal basis of $E_D \otimes_{\mathbf{Z}} \mathbf{R}$.

Then $\widehat{\deg}\,(\overline{E_D}) = 0$, and for any $n \in \mathcal{V}_D$, we have by the Poisson–Jensen formula (see for instance[29] [CDT21, § 2.4]),

$$h_\infty(\psi_D^{(n)}) \leq -n\log|\varphi'(0)| + D\int_{\mathbf{T}} \log^+|\varphi(z)|\,\mu_{\mathrm{Haar}}(z) + o(n+D)$$

and by the prime number theorem,

$$h_{\mathrm{fin}}(\psi_D^{(n)}) \leq \sigma_m n + o(n).$$

[29]We take $d = 1$ and $p(x) = x$ in [CDT21]. There is a minor difference with the assumptions in [CDT21]: we supposed there $f_i(\varphi(z))$ to be holomorphic on $|z| \leq 1$, whereas here we only assume that $f_i(\varphi(z))$ are meromorphic on $|z| < 1$. For all our estimates on archimedean heights in § 7, this difference is insignificant up to the error term of $o(n+D)$. The point is that for any $\varepsilon > 0$, we have $f_i(\varphi((1-\varepsilon)z))$ meromorphic on $|z| \leq 1$. There are at most finitely many meromorphic poles of all $f_i(\varphi((1-\varepsilon)z))$ in some neighborhood of $\overline{\mathbf{D}}$, and we may take a polynomial $h(z) \in \mathbf{C}[z]$ with $h(0) = 1$ and such that $h(z)f_i(\varphi((1-\varepsilon)z))$ are all holomorphic on a neighborhood of $|z| \leq 1$. We observe that replacing all $f_i(\varphi(z))$ by $h(z)f_i(\varphi((1-\varepsilon)z))$ yields the same archimedean estimate once we let $\varepsilon \to 0$ at the end. See the proof of Lemma 7.4.1 for details.

Fix an $\epsilon > 0$ with $\log|\varphi'(0)| > \sigma_m + \epsilon$. Then there is an $N_0 \in \mathbf{N}$ such that, for all $n \geq N_0$, we have

$$h_\infty(\psi_D^{(n)}) \leq -n\log|\varphi'(0)| + D\int_{\mathbf{T}} \log^+|\varphi(z)|\,\mu_{\mathrm{Haar}}(z) + (\epsilon/2)n + o(D);$$

$$h_{\mathrm{fin}}(\psi_D^{(n)}) \leq (\sigma_m + \epsilon/2)n.$$

Hence, for all $n \in \mathbf{N}$,

$$h_\infty(\psi_D^{(n)}) \leq -n\log|\varphi'(0)| + D\int_{\mathbf{T}} \log^+|\varphi(z)|\,\mu_{\mathrm{Haar}}(z) + (\epsilon/2)n + o(D);$$

$$h_{\mathrm{fin}}(\psi_D^{(n)}) \leq (\sigma_m + \epsilon/2)n + o(D).$$

By the slopes inequality (7.2.14) and $\log|\varphi'(0)| > \sigma_m + \epsilon$, it ensues that

$$\begin{aligned}
0 = \widehat{\deg}\,(\overline{E_D}) &\leq \sum_{n=0}^{\infty} \mathrm{rank}(E_D^{(n)}/E_D^{(n+1)}) \cdot h(\psi_D^{(n)}) = \sum_{n\in\mathcal{V}_D} h(\psi_D^{(n)}) \\
&\leq -\left(\sum_{n\in\mathcal{V}_D} n\right)(\log|\varphi'(0)| - \sigma_m - \epsilon) + (m'+1)D^2\int_{\mathbf{T}} \log^+|\varphi(z)|\mu_{\mathrm{Haar}}(z) + o(D^2) \\
&\leq -\left(\sum_{n=0}^{(m'+1)D-1} n\right)(\log|\varphi'(0)| - \sigma_m - \epsilon) + (m'+1)D^2\int_{\mathbf{T}} \log^+|\varphi(z)|\mu_{\mathrm{Haar}}(z) + o(D^2) \\
&= -\binom{(m'+1)D}{2}(\log|\varphi'(0)| - \sigma_m - \epsilon) + (m'+1)D^2\int_{\mathbf{T}} \log^+|\varphi(z)|\mu_{\mathrm{Haar}}(z) + o(D^2)
\end{aligned} \tag{7.3.18}$$

Comparing the leading asymptotic order D^2 coefficients and then letting $\epsilon \to 0$, we have

$$m'+1 \leq \frac{2\int_{\mathbf{T}} \log^+|\varphi(z)|\mu_{\mathrm{Haar}}(z)}{\log|\varphi'(0)| - \sigma_m} < \infty,$$

contrary to our assumption that m' could be arbitrarily large. □

Example 7.3.19. For the case (see § 13)

$$\mathbf{b} := \begin{pmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}^{\mathrm{t}}$$

of relevance to the proof of Theorem A, we compute

$$\tau^\flat(\mathbf{b}) = (2+2) - \frac{2\cdot 1^2 + 2\cdot 3^2}{14^2} = \frac{191}{49}.$$

Example 7.3.20. For the case (see § 14.5)

$$\mathbf{b} := \begin{pmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}^{\mathrm{t}}$$

of relevance to the proof of Theorem C, we also compute

$$\tau^\flat(\mathbf{b}) = (2+2) - \frac{2\cdot 1^2 + 2\cdot 3^2}{17^2} = \frac{1136}{289}.$$

7.4. **The convexity enhancement of the Bost–Charles bound.** We follow the same outline as in § 7.3, working with the same evaluation module $\overline{E}_D$ and using the same non-archimedean evaluation heights estimate. The improvement is from the optimal use of the dilated maps $\varphi_r(z) := \varphi(rz)$ at the step of the archimedean evaluation height estimate.

In order to estimate $\psi_D^{(n)}$ at the vanishing filtration jumps $n \in \mathcal{V}_D$, we consider $Q_i \in \mathbf{Z}[x^{-1}]_{\leq D} = \Gamma(\mathbf{P}^1_{\mathbf{Z}}, \mathcal{O}(D))$ such that $s := \sum_{i=1}^m f_i Q_i = c_n x^n + \dots$ has exact vanishing order n at $x = 0$. We can view s as a formal section of $\mathcal{O}(D)$, and then $s(x) \cdot x^D$ is canonically a formal function. Recall that we have endowed E_D with a Euclidean norm induced by the Hermitian line bundle $\overline{\mathcal{L}} = \overline{\mathcal{O}(1)}$, on which the Hermitian metric is induced by $\varphi_* \log^+ |z|^{-1}$. By extension, we define $\overline{\mathcal{L}}_r$ to be the line bundle $\mathcal{O}(1)$ equipped with the Hermitian metric induced from $(\varphi_r)_* \log^+ |z|^{-1}$. Explicitly:

$$\|\mathbf{1}(y)\|_{\overline{\mathcal{L}}_r} := \exp\left(- \sum_{z \in \varphi_r^{-1}(y)} \log^+ \frac{1}{|z|} \right) = \prod_{z \in D(0,r),\, \varphi(z) = y} |z/r|.$$

As a generalization of (7.3.9), we have the following archimedean evaluation height estimate in which we can take the optimal radius parameter $r = r(n)$:

Lemma 7.4.1. *For any $0 < r \leq 1$, we have*

$$h_\infty(\psi_D^{(n)}) \leq -n \log |\varphi_r'(0)| + D(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}_r) + o(D).$$

Proof. We assumed the functions $\varphi^* f_i(z) = f_i(\varphi(z))$ to be meromorphic on $|z| < 1$. Let us firstly remark that we can reduce the proof to the stronger assumption that $\varphi^* f_i \in \mathcal{M}(\overline{\mathbf{D}})$, namely that $f_i(\varphi(z))$ is meromorphic on an open neighborhood of $|z| \leq 1$, for all $1 \leq i \leq m$. Indeed, in the following proof, for $r < 1$, we only use the assumption that the $\varphi^* f_i$ are meromorphic on an open neighborhood of $|z| \leq r$; therefore we only need to discuss the reduction step for $r = 1$. Of course this particular $r = 1$ case is indeed the estimate by Bost and Charles recalled in (7.3.9). Nevertheless, we spell out a limit argument for deducing the $r = 1$ case from the $r < 1$ case, for the same reduction can be applied in the proofs in §§ 7–8 to allow us to assume $\varphi^* f_i \in \mathcal{M}(\overline{\mathbf{D}})$. Note that

$$\lim_{r \to 1} \log |\varphi_r'(0)| = \log |\varphi'(0)|, \qquad \lim_{r \to 1} \overline{\mathcal{L}} \cdot \overline{\mathcal{L}}_r = \overline{\mathcal{L}} \cdot \overline{\mathcal{L}}.$$

Therefore, the $r \to 1^-$ limit of the inequality on $h_\infty(\psi_D^{(n)})$ gives

$$h_\infty(\psi_D^{(n)}) \leq -n \log |\varphi'(0)| + D(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}) + o(n + D).$$

This gives the desired inequality with $r = 1$ since $o(n + D) = o(D)$ by Theorem 3.2.13 and Lemma 7.3.17.

And so we start from the meromorphy $\varphi^* f_i \in \mathcal{M}(\overline{\mathbf{D}})$ of all pullbacks. Choose and fix a holomorphic function $h \in \mathcal{O}(\overline{\mathbf{D}})$ such that $h(0) = 1$ and all $h \cdot \varphi^* f_i \in \mathcal{O}(\overline{\mathbf{D}})$ are holomorphic. We follow the notation $h_r(z) := h(rz)$ for $r \in (0, 1]$ and $z \in \overline{\mathbf{D}}$. Then, for any $s = \sum_{i=1}^m f_i Q_i = c_n x^n + \dots$ as above, $z^{-n} h_r(z) \cdot \varphi_r^* \left(s(x) \cdot x^D \right) \in \mathcal{O}(\overline{\mathbf{D}})$ is a holomorphic function whose $z = 0$ value equals $c_n \varphi_r'(0)^n \neq 0$. Therefore $\log \left| z^{-n} h_r(z) \cdot \varphi_r^* \left(s(x) \cdot x^D \right) \right|$ is a subharmonic function on $\overline{\mathbf{D}}$.

We modify the computation in [Bos20, §10.5.5]. Instead of using φ as in *loc. cit.*, we apply the Poisson–Jensen formula — or the subharmonic property — to this subharmonic function $\log|z^{-n}h_r\varphi_r^*(s(x)\cdot x^D)|$. This gets us the upper bound

$$\begin{aligned}\log|c_n| &\leq -n\log|\varphi_r'(0)| + \int_{\mathbf{T}} \log|h_r\varphi_r^*(s(x)\cdot x^D)|\,\mu_{\mathrm{Haar}} \\ &= -n\log|\varphi_r'(0)| + \int_{\mathbf{T}} \log|\varphi_r^*(s(x)\cdot x^D)|\,\mu_{\mathrm{Haar}} + O(1).\end{aligned} \tag{7.4.2}$$

Now we claim the identity

$$D(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r) = -\int_{\mathbf{T}} \log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}}\mu_{\mathrm{Haar}}. \tag{7.4.3}$$

To prove it, we start from the Poincaré–Lelong formula that gives:

$$\frac{i}{\pi}\partial\overline{\partial}\log\|\varphi_r^*(x^{-D})\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}} = -c_1\left(\varphi_r^*\overline{\mathcal{L}}^{\otimes D}\right),$$

$$\frac{i}{\pi}\partial\overline{\partial}\log^+|z|^{-1} = -\delta_0 + \mu_{\mathrm{Haar}}.$$

Therefore, by the Green–Stokes formula, we find

$$\begin{aligned}&-\int_{\mathbf{T}} \log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}}\mu_{\mathrm{Haar}} \\ &= \int_{\overline{\mathbf{D}}} -\log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}}\frac{i}{\pi}\partial\overline{\partial}\log^+|z|^{-1} - \log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}}|_{z=0} \\ &= \int_{\overline{\mathbf{D}}} -\frac{i}{\pi}\partial\overline{\partial}\log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}}\log^+|z|^{-1} + \widehat{\deg}\left(\varphi_r^*\overline{\mathcal{L}}^{\otimes D}|_{z=0}\right) \\ &= \int_{\overline{\mathbf{D}}} \log^+|z|^{-1}c_1(\varphi_r^*\overline{\mathcal{L}}^{\otimes D}) + \widehat{\deg}\left(\overline{\mathcal{L}}^{\otimes D}|_{x=0}\right) \\ &= D\cdot\left((\iota,\varphi_r)^*\overline{\mathcal{L}}\cdot([0],\log^+|z|^{-1})\right) \\ &= D\cdot\left(\overline{\mathcal{L}}\cdot(\iota,\varphi_r)_*([0],\log^+|z|^{-1})\right) \\ &= D\cdot\left(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r\right).\end{aligned}$$

proving (7.4.3).

At this point, by (7.4.2) and the pointwise decomposition

$$\log|\varphi_r^*(s(x)\cdot x^D)| = \log\|\varphi_r^* s\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}} - \log\|\varphi_r^* x^{-D}\|_{\varphi_r^*\overline{\mathcal{L}}^{\otimes D}} \tag{7.4.4}$$

inside the $\mathbf{T}$ integrands, the lemma follows if we prove an $o(D)$ bound on the $\mathbf{T}$ integral of the first term on the right-hand side of (7.4.4) under the assumption $\|s\|\leq 1$.

If in place of the integration measure μ_{Haar} we had a continuous measure μ on $\overline{\mathbf{D}}$, we would have had a constant C such that $C\varphi^*\nu \geq \mu$, and then we would have had

$$\begin{aligned}0 \geq \log\|s\| &\geq \max_{1\leq i\leq m}\log\|Q_i\| = \frac{1}{2}\max_{1\leq i\leq m}\log\int_{\mathcal{X}(\mathbf{C})}\|Q_i\|_{\mathcal{L}}^2\,\nu \\ &\geq C' + \frac{1}{2}\log\int_{\overline{\mathbf{D}}}\|\varphi^* s\|_{\varphi^*\overline{\mathcal{L}}}^2\,\mu \geq C' + \int_{\mathbf{T}}\log\|\varphi^* s\|_{\varphi^*\overline{\mathcal{L}}}\,\mu,\end{aligned}$$

where C' is a constant depending only on m, f_i, C, and is independent of D. We obtain the desired bound with μ_{Haar} in place of μ upon approximating $\log^+\frac{1}{|z|}$ (the Green function of μ_{Haar}) by smooth Green functions. See § 8.2.11 for details. □

Next, we explicitly calculate the arithmetic intersection number $\left(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r\right)$:

Lemma 7.4.5. *For* $0<r\le 1$, *we have*

$$\left(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r\right)=\int_{\mathbf{T}^2}\log|\varphi(z)-\varphi(rw)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w).$$

Proof. Recall from the discussion above that we have, straight from the definition,

$$\left(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r\right)=\int_{\overline{\mathbf{D}}}\log^+|z|^{-1}c_1(\varphi_r^*\overline{\mathcal{L}})+\widehat{\deg}\,(\overline{\mathcal{L}}|_{x=0}).$$

From the definition of $\overline{\mathcal{L}}$ and the functorial behavior of the Chern form under pushforward [BC22, Proposition 3.4.5(2)], we have

$$c_1(\varphi_r^*\overline{\mathcal{L}})=\varphi_r^*\varphi_*\mu_{\mathrm{Haar}},\qquad \left(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r\right)=\int_{\overline{\mathbf{D}}}\log^+|z|^{-1}\varphi_r^*\varphi_*\mu_{\mathrm{Haar}}+\widehat{\deg}\,(\overline{\mathcal{L}}|_{x=0}).$$

Recall again from the discussion above, and from $\|\mathbf{1}(x)\|_{\overline{\mathcal{L}}}=\prod_{z\in\overline{\mathbf{D}},\varphi(z)=x}|z|$ and using the Poisson–Jensen formula, that

$$\begin{aligned}\widehat{\deg}\,(\overline{\mathcal{L}}|_{x=0})&=-\log\|x^{-1}\|_{\overline{\mathcal{L}}}\big|_{x=0}=-\log\left((\varphi(z))^{-1}\prod_{z\in\overline{\mathbf{D}},\,\varphi(z)=x}|z|\right)\Bigg|_{x=0}\\&=\log|\varphi'(0)|+\sum_{0\neq z\in\overline{\mathbf{D}},\,\varphi(z)=0}\log|z|^{-1}\\&=\int_{\mathbf{T}}\log|\varphi(z)|\mu_{\mathrm{Haar}}(z),\end{aligned}$$

where both the product and sum count with multiplicities.

We follow the same computation as in [BC22, §5.4 and Example 5.3.2.1]. Note that on $\mathbf{C}^2$ (with coordinates x,y), we have

$$\frac{i}{\pi}\partial\overline{\partial}\log|x-y|^{-1}=-\delta_{\Delta(\mathbf{C})},$$

where $\Delta(\mathbf{C})$ denotes the diagonal divisor on $\mathbf{C}^2$.

For every $z\in\mathbf{T}\subset\overline{\mathbf{D}}$, we have (here we also view $\varphi(z)$ as the constant function that maps all points on $\overline{\mathbf{D}}$ to $\varphi(z)$)

$$\begin{aligned}\varphi_r^*\varphi_*\delta_z&=\varphi_r^*\delta_{\varphi(z)}=(\varphi_r,\varphi(z))^*\delta_{\Delta(\mathbf{C})}\\&=(\varphi_r,\varphi(z))^*\frac{i}{\pi}\partial\overline{\partial}\log|x-y|=\frac{i}{\pi}\partial\overline{\partial}\log|\varphi_r(w)-\varphi(z)|.\end{aligned}$$

Therefore for a fixed z, by using the Green–Stokes formula (it is alright here even though the Green functions are not smooth like in [BC22]), we have

$$\begin{aligned}\int_{\overline{\mathbf{D}}}\log^+|w|^{-1}\varphi_r^*\varphi_*\delta_z&=\int_{\overline{\mathbf{D}}}\log^+|w|^{-1}\frac{i}{\pi}\partial\overline{\partial}\log|\varphi_r(w)-\varphi(z)|\\&=\int_{\overline{\mathbf{D}}}\left(\frac{i}{\pi}\partial\overline{\partial}\log^+|w|^{-1}\right)\cdot\log|\varphi_r(w)-\varphi(z)|\\&=\int_{\overline{\mathbf{D}}}(\mu_{\mathrm{Haar}}(w)-\delta_{w=0})\cdot\log|\varphi_r(w)-\varphi(z)|\\&=\int_{\mathbf{T}}\log|\varphi_r(w)-\varphi(z)|\mu_{\mathrm{Haar}}(w)-\log|\varphi(z)|.\end{aligned}$$

We now integrate over z on $\mathbf{T}$ and then we have

$$\int_{\overline{\mathbf{D}}} \log^+ |w|^{-1} \varphi_r^* \varphi_* \mu_{\mathrm{Haar}} = \int_{\mathbf{T}^2} \log |\varphi_r(w) - \varphi(z)| \mu_{\mathrm{Haar}}(z) \mu_{\mathrm{Haar}}(w) - \int_{\mathbf{T}} \log |\varphi(z)| \mu_{\mathrm{Haar}}(z).$$

We arrive at the claimed formula

$$\begin{aligned}(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}_r) &= \int_{\overline{\mathbf{D}}} \log^+ |z|^{-1} \varphi_r^* \varphi_* \mu_{\mathrm{Haar}} + \int_{\mathbf{T}} \log |\varphi(z)| \mu_{\mathrm{Haar}}(z) \\ &= \int_{\mathbf{T}^2} \log |\varphi_r(w) - \varphi(z)| \mu_{\mathrm{Haar}}(z) \mu_{\mathrm{Haar}}(w).\end{aligned}$$

□

Proof of Theorems 7.1.6 and 7.1.10. By Lemma 7.1.4, we have that $\widehat{T}(r, \varphi)$ as a function in $\log r$ is nondecreasing and convex. Therefore it suffices to prove Theorem 7.1.6. By Lemmas 7.4.1 and 7.4.5, using the slopes notation (7.1.7), we have that for $n/D \in [\alpha_k, \alpha_{k+1}]$, the optimal bound for $h_\infty(\psi_D^{(n)})$ is obtained by using r_k among $0 \le k \le l$; here we set $\alpha_0 = 0, \alpha_{l+1} = m$. Note, once again, that by Corollary 2.6.1 and Theorem 3.2.13 we have (for any $\varepsilon > 0$ and $D \gg_\varepsilon 1$) the containment $\mathcal{V}_D \subset [0, (m+\varepsilon)(D+1) + C(\varepsilon)]$. For $n \in [mD, (m+\varepsilon)(D+1) + C(\varepsilon)]$, we use the bound for $h_\infty(\psi_D^{(n)})$ obtain with the full radius $r_l = 1$. Letting $\varepsilon \to 0$, by a similar argument to the proof of Theorem 7.0.1 (see the proof of (7.3.15)), we deduce from Lemmas 7.4.1 and 7.4.5 and a straightforward computation that

$$\begin{aligned}&\sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)}) \\ &\le \left(-\frac{m^2}{2} \log |\varphi'(0)| + \sum_{k=0}^{l} (\alpha_{k+1} - \alpha_k) \widehat{T}(r_k, \varphi) - \frac{1}{2}(\alpha_{k+1}^2 - \alpha_k^2) \log r_k \right) D^2 + o(D^2) \\ &= \left(-\frac{m^2}{2} \log |\varphi'(0)| + m\widehat{T}(1, \varphi) - \frac{1}{2} \sum_{k=1}^{l} \alpha_k^2 (\log r_k - \log r_{k-1}) \right) D^2 + o(D^2).\end{aligned}$$

Then the desired bound follows this estimate combined with (7.3.15) and (7.3.7).

□

Example 7.4.6. In the proof of Theorem A, we use φ as in § A.5, where we have $(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}) = 11.844\ldots$ and $m = 14$.

We first apply Theorem 7.1.6 with $l = 1$ and $r_0 = e^{-1/2}, r_1 = 1$. We compute

$$(\overline{\mathcal{L}}_{e^{-1/2}} \cdot \overline{\mathcal{L}}) = 10.5739\ldots,$$

and the slope

$$\alpha_1 = \frac{(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}) - (\overline{\mathcal{L}}_{e^{-1/2}} \cdot \overline{\mathcal{L}})}{\log r_1 - \log r_0} = 2.5410\ldots.$$

Therefore, the convexity saving is

$$\frac{\frac{1}{m}\alpha_1^2(\log r_1 - \log r_0)}{\log |\varphi'(0)| - \tau(\mathbf{b}; \mathbf{e})} = 0.27243\ldots.$$

In other words, we obtain the proof by contradiction with

$$\frac{(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}) - \frac{1}{m}\alpha_1^2(\log r_1 - \log r_0)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} = 13.99303\ldots - 0.27243\ldots = 13.7206\ldots < 14. \tag{7.4.7}$$

We refine this further by taking more radii:

$$r_0 = e^{-1}, \quad r_1 = e^{-1/2}, \quad r_2 = e^{-1/4}, \quad r_3 = 1.$$

At these radii, we compute the corresponding Bost–Charles characteristic integrals:

$$\begin{aligned}
\widehat{T}(r_3,\varphi) &= (\overline{\mathcal{L}},\overline{\mathcal{L}}) = 11.844\ldots,\\
\widehat{T}(r_2,\varphi) &= (\overline{\mathcal{L}}_{e^{-1/4}}\cdot\overline{\mathcal{L}}) = 11.049\ldots,\\
\widehat{T}(r_1,\varphi) &= (\overline{\mathcal{L}}_{e^{-1/2}}\cdot\overline{\mathcal{L}}) = 10.573\ldots,\\
\widehat{T}(r_0,\varphi) &= (\overline{\mathcal{L}}_{e^{-1}}\cdot\overline{\mathcal{L}}) = 9.8766\ldots;
\end{aligned}$$

and the corresponding slopes:

$$\alpha_1 = 1.3943\ldots, \alpha_2 = 1.9018\ldots, \alpha_3 = 3.1802\ldots.$$

We thus derive the following convexity saving in the holonomy bound:

$$\frac{\frac{1}{m}\sum_{k=1}^{3}\alpha_k^2(\log r_k - \log r_{k-1})}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} = 0.37171\ldots;$$

In other words, the refined holonomy bound is

$$13.99303\ldots - 0.37171\ldots = 13.621\ldots \tag{7.4.8}$$

Example 7.4.9. In order to prove Theorem A, as discussed in Remark A.5.2, we could try to only use the 9 functions (without integrations). Then Theorem 7.1.6 with $l = 3$ in the above example gives

$$\frac{11.844\ldots - \frac{1}{9}\sum_{k=1}^{3}\alpha_k^2(\log r_k - \log r_{k-1})}{\log\left(256\cdot\frac{5448339453535586608000000000}{8658833407565631122430056127}\right) - 2\cdot 157/81} = 9.4203\ldots < 10,$$

which comes nearer to the 9 threshold but it remains insufficient to draw a contradiction with 9 functions. This is why we need the integrations idea.

7.5. **Binomial metrics: proof of Theorem 7.1.13.** We recall our assumption on the denominator types of $\{f_i\}$. Set $u_0 := 0$ and $u_{r+1} := m$. For $0 \le h \le r$, if $u_h < i \le u_{h+1}$, then

$$f_i(x) = a_{i,0} + \sum_{n=1}^{\infty} a_{i,n}\frac{x^n}{n^{e_i}[1,\ldots,b_1\cdot n]\cdots[1,\ldots,b_h\cdot n]}, \qquad a_{i,n}\in\mathbf{Z}.$$

We take our evaluation module to be the following free $\mathbf{Z}$-module of rank mD:

$$E_D = \bigoplus_{h=0}^{r}\bigoplus_{i=u_h+1}^{u_{h+1}}\frac{[1,\ldots,\xi D]^{e_i}}{[1,\ldots,u_{h+1}b_{h+1}D]\cdots[1,\ldots,u_r b_r D]}f_i\,\mathbf{Z}[x]_{<D}.$$

We endow E_D with the Euclidean norm that has $\{f_i x^k\}_{1\le i\le m, 0\le k<D}$ as an orthogonal basis with vector lengths $\|x^k\| = e^{D(\lambda t^r + \mu t)}$, where $t = k/D$. Recall from our assumption that $\lambda > 0$ and $r > 1$. We use $\overline{E}_D$ to denote this Euclidean lattice.

Applying the defining formula (7.2.3) of $\widehat{\deg}\,\overline{E}_D$ to the $\mathbf{Q}$-basis $\{f_i x^k\}_{1\le i\le m, 0\le k<D}$ of $E_D \otimes \mathbf{Q}$, we compute, as $D \to \infty$:

$$\begin{aligned}\widehat{\deg}\,\overline{E}_D &= -m\left(\int_0^1 \lambda t^r + \mu t\,dt\right)D^2 + \left(\sum_{h=0}^{r-1}(u_{h+1}-u_h)\sum_{j=h+1}^{r} u_j b_j\right)D^2 \\ &\quad - \left(\xi\sum_{i=1}^{m} e_i\right)D^2 + o(D^2) \\ &= -m\left(\frac{\lambda}{r+1}+\frac{\mu}{2}\right)D^2 + \left(\sum_{h=1}^{r} u_h^2 b_h - \xi\sum_{i=1}^{m} e_i\right)D^2 + o(D^2).\end{aligned} \tag{7.5.1}$$

Again we let $F_{\mathbf{Q}} := \mathbf{Q}[\![x]\!]$, and we filter it by the $x=0$ vanishing order:

$$F_{\mathbf{Q}} = F_{\mathbf{Q}}^{(0)} \supseteq F_{\mathbf{Q}}^{(1)} \supseteq \cdots \supseteq F_{\mathbf{Q}}^{(n)} \supseteq \cdots,$$

where

$$F_{\mathbf{Q}}^{(n)} := \mathrm{Span}_{\mathbf{Q}}\{x^k \,:\, k \ge n\}.$$

The graded piece $F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}$ is a one dimensional $\mathbf{Q}$-vector space generated by the image of x^n under the quotient map. The Euclidean lattice structure on $F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}$ is given by the free rank one $\mathbf{Z}$-module generated by the image of x^n and the Euclidean norm with $\|x^n\| = 1$. This is the same as in § 7.3 up to a shift by $-D$ in the power of x.

As in § 7.3, we the have natural injective *evaluation map* $\psi_D : E_D \to F_{\mathbf{Q}}$, inducing injections on the graded pieces $\psi_D^{(n)} : E_D^{(n)}/E_D^{(n+1)} \to F_{\mathbf{Q}}^{(n)}/F_{\mathbf{Q}}^{(n+1)}$. We still have $\mathrm{rank}\, E_D^{(n)}/E_D^{(n+1)} \in \{0,1\}$, and the cardinality $\#\mathcal{V}_D = \mathrm{rank}\, E_D = mD$ of the vanishing filtration jumps set $\mathcal{V}_D = \left\{n \in \mathbf{N} \,:\, \mathrm{rank}\, E_D^{(n)}/E_D^{(n+1)} = 1\right\}$.

We now provide upper bounds on the evaluation heights $h_\infty(\psi_D^{(n)})$ and $h_{\mathrm{fin}}(\psi_D^{(n)})$. For $v \in M_{\mathbf{Q}}$, by the definition of the local evaluation height h_v, we consider an arbitrary $(Q_i)_{1\le i\le m} \in E_D^{(n)} \smallsetminus E_D^{(n+1)}$, and our task is to provide an upper bound on $\log|c_n|_v - \log\|(Q_i)_{1\le i\le m}\|_{E_D,v}$, where c_n denotes the coefficient of x^n in $\sum_{i=1}^m f_i(x)Q_i(x)$, and $|\cdot|_v$ is the usual v-adic norm on $\mathbf{Q}$.

For $v = \infty$, we use the equivalent interpretation of $h_\infty(\psi_D^{(n)})$ by considering any $(Q_i)_{1\le i\le m} \in E_{D,\mathbf{R}}^{(n)} \smallsetminus E_{D,\mathbf{R}}^{(n+1)}$ with $\|(Q_i)_{1\le i\le m}\|_{E_D,\infty} \le 1$, and providing an upper bound on $\log|c_n|_\infty$, which is then our upper bound for $h_\infty(\psi_D^{(n)})$. By definition of our binomial metric, upon writing momentarily $t := k/D$, the unit ball condition $\|(Q_i)_{1\le i\le m}\|_{E_D,\infty} \le 1$ implies the bounds $|\alpha_{i,k}|_\infty \le e^{-D(\lambda t^r+\mu t)}$ on the coefficients of $Q_i(x) = \sum_{k=0}^{D-1}\alpha_{i,k}x^k$, for all $1 \le i \le m$.

For simplicity of notation, we write $f_i(x) = \sum_{n=0}^\infty a'_{i,n}x^n$. By assumption, all f_i converge on the closed disc $\overline{D(0,\rho)}$ for all $\rho < R$. We use this information to derive an upper bound on $h_\infty(\psi_D^{(n)})$ which is useful on a certain range of n/D. The analyticity on $\overline{\mathbf{D}}_\rho$ means that $|a'_{i,k}|_\infty = O_\rho(\rho^{-k})$, where the implicit constant only depends on $\rho, f_1, \dots, f_m$, but not on k. Hence, for arbitrary $n \in \mathbf{N}$, we derive from

$$c_n = \sum_{i=1}^{m}\sum_{k=0}^{\min(n,D-1)} \alpha_{i,k}a'_{i,n-k}, \text{ thus } |c_n|_\infty \le mD \max_{\substack{1\le i\le m,\\ 0\le k\le \min(n,D-1)}} |\alpha_{i,k}|_\infty \cdot |a'_{i,n-k}|_\infty.$$

the following archimedean evaluation heights bound:

$$h_\infty(\psi_D^{(n)}) \leq \left(\max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log\rho\} \right) D + o_\rho(D). \quad (7.5.2)$$

The function of $t \in [0, \min\{1, n/D\}]$ under the maximum is concave, and in particular unimodal. From here it is easy to justify the $\rho \to R^-$ limit:

$$h_\infty(\psi_D^{(n)}) \leq \left(\max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log R\} \right) D + o(n+D). \quad (7.5.3)$$

We include the details of this limiting argument as it also reveals the limit point $\rho = R^-$ to indeed be the optimal choice to make in (7.5.2) across $\rho \in (0, R)$. Let us denote the maximizers of the unimodal functions under the curly brackets in (7.5.3) and (7.5.2) to be at $t := t_R$ and $t := t_\rho$, respectively. We have, noting that by definition $0 \leq t_R, t_\rho \leq \min\{1, n/D\}$:

$$\begin{aligned}
&\max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log\rho\} \geq -\lambda t_R^r - \mu t_R - (n/D - t_R)\log\rho \\
&= (\log R - \log\rho)(n/D - t_R) + \max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log R\} \\
&\geq \max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log R)\},
\end{aligned}$$

and similarly,

$$\begin{aligned}
&\max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log R\} \geq -\lambda t_\rho^r - \mu t_\rho - (n/D - t_\rho)\log R \\
&= -(\log R - \log\rho)(n/D - t_\rho) + \max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log\rho\} \\
&\geq -(\log R - \log\rho)(n/D) + \max_{0\leq t\leq \min\{1,n/D\}} \{-\lambda t^r - \mu t - (n/D - t)\log\rho\}.
\end{aligned}$$

This proves (7.5.3), and also the optimality of taking the limit $\rho \to R^-$ in (7.5.2).
Continuing with the proof, we set $s := n/D$, and recall our notation

$$\chi_0 := \min\left\{1, \left(\frac{\max\{0, \log R - \mu\}}{\lambda r}\right)^{1/(r-1)}\right\}$$

from the statement of the theorem under proof. Its meaning is the following. By the same computation as in the proof of Lemma 7.1.15, the following archimedean evaluation height bound is valid in the range $s \geq \chi_0$:

$$h_\infty(\psi_D^{(n)}) \leq \big(\Gamma(\log R, r, \lambda, \mu) - s\log R\big)D + o(n+D); \quad (7.5.4)$$

whereas in the range $0 \leq s \leq \chi_0$, the following improvement holds:

$$h_\infty(\psi_D^{(n)}) \leq (-\lambda s^r - \mu s)D + o(n) + o(D). \quad (7.5.5)$$

Saying that $s \in [0, \chi_0]$ is the domain of improvement of (7.5.5) over (7.5.4) is exactly the definition of χ_0.

For the range $s > \chi_0$, we instead use the Poisson–Jensen formula applied to the logarithm of the holomorphic function

$$h(z) \cdot \varphi^* \left(\sum_{i=1}^m f_i Q_i \right) \cdot z^{-n} \in \mathcal{O}(\overline{\mathbf{D}}),$$

where an arbitrary $h \in \mathcal{O}(\overline{\mathbf{D}})$ is fixed subject to $h(0) = 1$ and $h \cdot \varphi^* f_i \in \mathcal{O}(\overline{\mathbf{D}})$ for all $i = 1, \ldots, m$. We derive the usual bound (here we also use $|\cdot|_\infty$ to denote the usual absolute value on $\mathbf{C}$):

$$\begin{aligned}
&\log |c_n|_\infty \\
&\leq -n \log |\varphi'(0)| + \int_{\mathbf{T}} \left| h \cdot \varphi^* \left(\sum_{i=1}^m f_i Q_i \right) \right|_\infty \mu_{\mathrm{Haar}} \\
&\leq -n \log |\varphi'(0)| + \int_{\mathbf{T}} \max_{1 \leq i \leq m, 0 \leq k \leq \min\{n, D-1\}} (\log |\alpha_{i,k}|_\infty + k \log |\varphi(z)|_\infty) \, \mu_{\mathrm{Haar}} + o(D) \\
&\leq -n \log |\varphi'(0)| + \int_{\mathbf{T}} \max_{0 \leq k \leq \min\{n, D-1\}} (D(-\lambda (k/D)^r - \mu(k/D)) + k \log |\varphi(z)|_\infty) \, \mu_{\mathrm{Haar}} + o(D).
\end{aligned}$$

Therefore, by the definition of $T(\varphi; r, \lambda, \mu)$ via the Legendre transform $\Gamma(x; r, \lambda, \mu)$ of the binomial metric weight function $\lambda t^r + \mu t$, we derive the following for our upper bound on all the archimedean evaluation heights:

$$h_\infty(\psi_D^{(n)}) \leq -n \log |\varphi'(0)| + DT(\varphi; r, \lambda, \mu) + o(D). \tag{7.5.6}$$

Note that bound (7.5.4) is better than (7.5.6) if and only if

$$\frac{n}{D} \leq \chi_1 := \frac{T(\varphi; r, \lambda, \mu) - \Gamma(\log R; r, \lambda, \mu)}{\log |\varphi'(0)| - \log R}.$$

Hence, by Theorem 3.2.13 (letting $\varepsilon \to 0$ right after taking $D \to \infty$), we have

$$\begin{aligned}
&\limsup_{D \to \infty} \left\{ D^{-2} \sum_{n \in \mathcal{V}_D} h_\infty(\psi_D^{(n)}) \right\} \\
&\leq \int_0^{\chi_0} (-\lambda s^r - \mu s) \, ds + \int_{\chi_0}^{\chi_1} \big(\Gamma(\log R; r, \lambda, \mu) - s \log R \big) \, ds \\
&\quad + \int_{\chi_1}^m \big(T(\varphi; r, \lambda, \mu) - s \log |\varphi'(0)| \big) \, ds \\
&= mT(\varphi; r, \lambda, \mu) - \frac{m^2}{2} \log |\varphi'(0)| - \frac{(T(\varphi; r, \lambda, \mu) - \Gamma(\log R; r, \lambda, \mu))^2}{2(\log |\varphi'(0)| - \log R)} \\
&\quad - \chi_0 \Gamma(\log R; r, \lambda, \mu) + \chi_0^2 (\log R - \mu) \left(\frac{1}{2} - \frac{1}{r(r+1)} \right).
\end{aligned} \tag{7.5.7}$$

Next we turn to estimating $h_{\mathrm{fin}}(\psi_D^{(n)})$. Considering an arbitrary $(Q_i)_{1 \leq i \leq m} \in E_D^{(n)} \setminus E_D^{(n+1)}$, our task is for each prime p to provide an upper bound on $\log |c_n|_p$. Since the $\mathbf{Z}$-lattice E_D here has essentially the same structure as the one in § 7.3, the argument there yields for the total finite evaluation height the upper bound:

$$\limsup_{D \to \infty} \left\{ D^{-2} \sum_{n \in \mathcal{V}_D} h_{\mathrm{fin}}(\psi_D^{(n)}) \right\} \leq \frac{1}{2} \left(\sigma_m m^2 + \sum_{h=1}^r u_h^2 b_h \right) + \left(\max_{1 \leq i \leq m} e_i \right) I_\xi^m(\xi). \tag{7.5.8}$$

We plug (7.5.1), (7.5.7), and (7.5.8) into (7.2.14) and derive:

$$\frac{m^2}{2}\left(\log|\varphi'(0)| - \sigma_m + \frac{1}{m^2}\left(\sum_{h=1}^{r} u_h^2 b_h - \frac{2}{m^2}(\xi\sum_{i=1}^{m} e_i + \left(\max_{1\le i\le m} e_i\right) I_\xi^m(\xi))\right)\right)$$
$$\le m\left(T(\varphi;r,\lambda,\mu) + \frac{\lambda}{r+1} + \frac{\mu}{2}\right) - \frac{(T(\varphi;r,\lambda,\mu) - \Gamma(\log R;r,\lambda,\mu))^2}{2(\log|\varphi'(0)| - \log R)}$$
$$- \chi_0\Gamma(\log R;r,\lambda,\mu) + \chi_0^2(\log R - \mu)\left(\frac{1}{2} - \frac{1}{r(r+1)}\right),$$

which rearranges into the claimed bound on m. □

Example 7.5.9. In the proof of Theorem A, we use φ as in § A.5, and recall that

$$m = 14,$$
$$\log|\varphi'(0)| = \log\left(256\cdot\frac{5448339453535586608000000000}{8658833407565631122430056127}\right),$$
$$\tau(\mathbf{b};\mathbf{e}) = \frac{27}{80} + \frac{191}{49},$$

and that all f_i have convergence radii at least $R := 4$.

Select the following for the binomial metric weight parameters:

$$r = 4.7, \lambda = 10, \mu = -4.5.$$

A numerical computation gives

$$T(\varphi;4.7,10,-4.5) = 6.5316\ldots,$$
$$\Gamma(\log 4;4.7,10,-4.5) = 2.6429\ldots,$$

with

$$\chi_1 = 1.0522\ldots,\quad \chi_0 = 0.57035\ldots,$$

meeting the special assumptions that we made in Theorem 7.1.13, and supplying the holonomy bound $m \le 13.8527\ldots < 14$. The contradiction supplies a proof of Theorem A (see § 13), with a better numeric than when we use Theorem 7.0.1 alone like in § A.5, prior to the convexity enhancement by Theorem 7.1.10.

If we only work with 9 functions as in Remark A.5.2, with the same parameters above replacing $m = 9, \tau(\mathbf{b}';0) = 2\cdot\frac{157}{81}$, we have the bound in Theorem 7.1.13 is $9.5234\ldots < 10$, but not enough to draw a contradiction to deduce Theorem A. △

7.6. **A further improvement.** The setup is similar to Theorem 7.1.6: fix a set of subradii $1 = r_l > r_{l-1} > \cdots > r_0 > 0$. The following is the counterpart — and ultimate sharpening — of the archimedean term in the refined bound from § 6. In place of using the Hermitian line bundle $\overline{\mathcal{L}} = (\iota,\varphi)_*([0],\log^+|z|^{-1})$, we introduce weights $s_0,\ldots,s_l \in [0,1]$ with total mass $\sum_{h=0}^{l} s_h = 1$, and use the $\mathbf{s}$-weighted average of the Hermitian line bundles defined by the restricted maps $\varphi(r_k z)$:

$$\overline{\mathcal{L}}' := \prod_{h=0}^{l} \overline{\mathcal{L}}_{r_h}^{\otimes s_h}.$$

Here, by a mild abuse of **R**-line bundle notation, this is the line bundle $\mathcal{L} = \mathcal{O}(1)$ over $\mathcal{X}$ with Hermitian metric defined by

$$\|\cdot\|_{\overline{\mathcal{L}}'} := \prod_{h=0}^{l} \|\cdot\|_{\overline{\mathcal{L}}_{r_h}}^{s_h}.$$

As in § 7.4, for $1 \leq k \leq l$, set

$$\beta_k := \frac{\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_{k-1}}}{\log r_k - \log r_{k-1}} = \frac{\sum_{h=0}^{l} s_h(\overline{\mathcal{L}}_{r_h} \cdot \overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}_{r_h} \cdot \overline{\mathcal{L}}_{r_{k-1}})}{\log r_k - \log r_{k-1}}. \tag{7.6.1}$$

We note that by the same argument[30] as in Lemma 7.4.5 we have

$$\overline{\mathcal{L}}_{r_h} \cdot \overline{\mathcal{L}}_{r_k} = \int_{\mathbf{T}^2} \log|\varphi(r_h z) - \varphi(r_k w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w).$$

Therefore, since all $s_h \geq 0$, Lemma 7.1.4 on convexity shows that

$$0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_l.$$

As in Theorem 7.1.6, we assume $\beta_l \leq m$. (If this condition fails, it usually serves as a stronger bound on m anyhow.) We extend the notation by setting $\beta_0 := 0$ and $\beta_{l+1} := m$. For $n/D \in [\beta_k, \beta_{k+1})$, we estimate the archimedean evaluation height in terms of φ_{r_k}. Namely, the proof of Lemma 7.4.1 with $\overline{\mathcal{L}}$ replaced by $\overline{\mathcal{L}}'$ gives

$$h_\infty(\psi_D^{(n)}) \leq -n\log|\varphi'(0)| - n\log r_k + D(\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_k}) + o(D).$$

Similarly to the proof of Theorem 7.1.6, using Theorem 3.2.13, we have that $\mathcal{V}_D \subset [0, (m+\epsilon)D]$ once $D \gg_\epsilon 1$, and for $n \in [mD, (m+\epsilon)D]$, we continue to use the bound on $h_\infty(\psi_D^{(n)})$ from taking the full radius $r_l = 1$. As $D \to \infty$ and then $\epsilon \to 0$, we obtain

$$\limsup_{D\to\infty} \left\{ D^{-2} \sum_{n\in\mathcal{V}_D} h_\infty(\psi_D^{(n)}) \right\} \leq -\frac{m^2}{2}\log|\varphi'(0)| + \sum_{k=0}^{l} (\beta_{k+1} - \beta_k)\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_k} - \frac{1}{2}(\beta_{k+1}^2 - \beta_k^2)\log r_k.$$

We have the same estimate on $h_{\mathrm{fin}}(\psi_D^{(n)})$ as in (7.3.15). Finally,

$$\widehat{\deg}\, \overline{E}_D = \left(\frac{m}{2}(\overline{\mathcal{L}}'.\overline{\mathcal{L}}') + \sum_{h=1}^{r} u_h y_h - \xi\Big(\sum_{i=1}^{m} e_i\Big) \right) D^2 + o(D^2).$$

[30] These are actually the same statement upon changing $\varphi(z)$ to $\varphi(r_k z)$ and r to r_h/r_k, if $h \leq k$.

Hence, by the slopes inequality (7.2.14) as before, we get

$$\begin{aligned}
&m^2\left(\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})\right)\\
&\leq -m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' + \sum_{k=0}^{l}\Big(-(\beta_{k+1}^2-\beta_k^2)\log r_k + 2(\beta_{k+1}-\beta_k)\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_k}\Big)\\
&= 2m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1 - m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' - \sum_{k=1}^{l}\Big(-\beta_k^2(\log r_k - \log r_{k-1}) + 2\beta_k(\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_{k-1}})\Big)\\
&= 2m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1 - m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' - \sum_{k=1}^{l}\frac{(\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_{k-1}})^2}{\log r_k - \log r_{k-1}}\\
&= 2m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1 - m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' - \sum_{k=1}^{l}\beta_k(\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_{k-1}})\\
&= m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1 - m\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' + \sum_{k=0}^{l}(\beta_{k+1}-\beta_k)\overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_{r_k}.
\end{aligned} \tag{7.6.2}$$

Note that (7.6.2) gives a bound on m for every choice of partition $\mathbf{s} = \{s_h\}_{h=0}^l$, and it recovers the convexity saving of Theorem 7.1.6 as the special case $\mathbf{s} = (0,0,\ldots,0;1)$.

We propose the following choice of $\{s_h\}_{h=0}^l$. Let us postulate the following system of $l+1$ inhomogeneous linear equations in the $l+1$ unknowns s_h:

$$s_h = \frac{1}{m}(\beta_{h+1}-\beta_h), \quad 0\leq h\leq l, \quad \beta_0 = 0,\ \beta_{l+1} = m. \tag{7.6.3}$$

This choice is explained in Remark 7.6.7 below. We suppose the $(l+1)\times(l+1)$ coefficient matrix of this linear system to have a nonzero determinant, and furthermore that the unique solution $\{s_h^*\}$ has nonnegative components. This solution clearly has $\sum_{h=0}^l s_h^* = 1$, and we can set our $s_h := s_h^*$ in (7.6.2). The inequality (7.6.2) then reads:

$$m^2(\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})) \leq m\sum_{h=0}^{l} s_h^*\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}_1.$$

In this situation we derive the following refined holonomy bound:

$$m \leq \frac{\sum_{h=0}^l s_h^*\cdot\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}_1}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}.$$

We summarize our findings into a theorem:

Theorem 7.6.4. *Assume the same conditions and notation as in Theorem 7.0.1. Fix a sequence of subradii $1 = r_l > r_{l-1} > \cdots > r_0 > 0$. Assume that the following system of $l+1$ linear inhomogeneous equations in the $l+1$ unknowns $\{s_h\}_{h=0}^l$*

$$\begin{aligned}
&m\sum_{h=0}^{k-1} s_h = \frac{\sum_{h=0}^l s_h(\overline{\mathcal{L}}_{r_h}\cdot\overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}_{r_h}\cdot\overline{\mathcal{L}}_{r_{k-1}})}{\log r_k - \log r_{k-1}}, \qquad k = 1,\ldots,l,\\
&\sum_{h=0}^{l} s_h = 1
\end{aligned} \tag{7.6.5}$$

has a unique solution $\{s_h^*\}\in[0,1]^{l+1}$.

Then,

$$m \leq \frac{\sum_{h=0}^{l} s_h^* \overline{\mathcal{L}}_{r_k} \cdot \overline{\mathcal{L}}_1}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} = \frac{\overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_1 - \sum_{h=0}^{l-1} s_h^* (\overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_1 - \overline{\mathcal{L}}_{r_h} \cdot \overline{\mathcal{L}}_1)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}. \tag{7.6.6}$$

Remark 7.6.7. The special assumptions about the linear system (7.6.5) having a unique solution with nonnegative components appears to hold in practice. We do not know if it is a general feature. The heuristic behind this particular choice of $s_h = s_h^*$ is to emulate the Euler–Lagrange stationary action principle on our upper bound (7.6.2). Namely, we compute the d/ds_h derivatives of that upper bound,

$$2m\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_1 - m\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}' - \sum_{k=1}^{l} \frac{(\overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_k} - \overline{\mathcal{L}}' \cdot \overline{\mathcal{L}}_{r_{k-1}})^2}{\log r_k - \log r_{k-1}}$$

and set these derivatives to 0. △

Example 7.6.8. Let us revisit now the first case in Example 7.4.6: $l = 1, r_0 = e^{-1/2}$, and 14 putative functions for Theorem A. We have

$$C := \overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_1 = 11.844\ldots$$
$$B := \overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_{e^{-1/2}} = 10.573\ldots$$
$$A := \overline{\mathcal{L}}_{e^{-1/2}} \cdot \overline{\mathcal{L}}_{e^{-1/2}} = 8.3717\ldots.$$

Now the point of the improvement over the previous bound is that

$$2\overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_{e^{-1/2}} > \overline{\mathcal{L}}_1 \cdot \overline{\mathcal{L}}_1 + \overline{\mathcal{L}}_{e^{-1/2}} \cdot \overline{\mathcal{L}}_{e^{-1/2}}.$$

We calculate

$$\beta_1 = \frac{s_0(B-A) + s_1(C-B)}{\log r_1 - \log r_0} = 2(s_0(B-A) + s_1(C-B)) > 2(C-B) = \alpha_1,$$

whence

$$s_0^* = \frac{1}{m}\beta_1(s_0^*, s_1^*) > \alpha_1/m.$$

The new convexity saving is by $\frac{s_0^*(C-B)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} > \frac{\alpha_1(C-B)/m}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})}$, where the latter was the previous convexity saving in Example 7.4.6. Explicitly, we find s_0^* and s_1^* as the solution of the two linear equations

$$s_0 = \frac{2}{m}(s_0(B-A) + s_1(C-B)), \quad s_1 = 1 - \frac{2}{m}(s_0(B-A) + s_1(C-B)).$$

The solution is

$$s_0^* = \frac{C-B}{m/2 + (A+C-2B)} = 0.20936\ldots$$

resulting in the following improvement convexity saving over Example 7.4.6 (compare with equation (7.4.7)):

$$\frac{s_0^*(C-B)}{\log|\varphi'(0)| - \tau(\mathbf{b};\mathbf{e})} = 0.31426\ldots(> 0.27243\ldots).$$

In other words, the refined holonomy bound on m is here

$$13.99303\ldots - 0.31426\ldots = 13.678\ldots, \tag{7.6.9}$$

giving a still more comfortable numerical margin for the ultimate contradiction to $m = 14$. (A similar but more complicated computation with four radii instead of two would also yield a slight improvement on equation (7.4.8) in Example 7.4.6.) △

At this point, a reader primarily interested in the proof of Theorems A and C can skip directly ahead to § 9 on a first reading.

7.7. **On bypassing the Kolchin–Shidlovsky type theorems from § 3.2.** At least for our specific and qualitative applications in the present paper, it is technically possible to avoid all recourse to the — fairly technical — zero estimates we collected in § 3.2. As the general case of our abstract holonomy bounds seems somewhat awkward to approach in its full generality[31] without using the functional bad approximability theorems, while on the other hand the Shidlovsky (or Chudnovsky–Osgood) type of input is a golden standard in the subject which — furthermore and more importantly — turns out indispensable for all quantitative refinements in our method to deriving actual Diophantine inequalities on the bad approximability of a period vector by an integer vector, we limit ourselves here to only a few brief indications on how one could technically avoid the appeal to § 3.2 or the purpose of proving certain relaxed versions of our holonomy bounds, still sufficient for all our present applications in this paper.

We recall that for all the proofs in this section, we have made the assumption that $0 = u_0 \le u_1 \le \cdots \le u_r < u_{r+1} = m$ in the denominators form (6.0.3), as permutation on the columns of $\mathbf{b}$ does not change the assumption on f_i.

7.7.1. *Discussion for Theorem 7.0.1.* We sketch a proof of (7.0.3) that bypasses § 3.2 under assuming the stronger positivity condition $\log|\varphi'(0)| > \sigma_m + \max(e_i)$ in place of (7.0.2). In our application to Theorem A, and to at least some weaker form (i.e., with 10^{-6} replaced by a smaller explicit positive number) of Theorem C, this condition is satisfied since $\log|\varphi'(0)| = \log\left(256 \cdot \frac{5448339453535586608000000000}{8658833407565631122430056127}\right) > 5.08 > 4 + 1 = \sigma_m + \max_{i=1}^m(e_i)$ in § A.5.

Let $\overline{h}_\infty(\psi_D^{(n)}), \overline{h}_{\mathrm{fin}}(\psi_D^{(n)})$ denote the main terms in the bounds on the archimedean, resp. finite evaluation heights $h_\infty(\psi_D^{(n)}), h_{\mathrm{fin}}(\psi_D^{(n)})$ that we proved in (7.3.9), resp. (7.3.11), (7.3.12), and (7.3.13); in $\overline{h}_{\mathrm{fin}}$, we take the optimal parameters choices $y_h := b_h u_h$ that we used at the end of our proof. Thus the global evaluation height has an upper bound with main term $\overline{h}(\psi_D^{(n)}) = \overline{h}_\infty(\psi_D^{(n)}) + \overline{h}_{\mathrm{fin}}(\psi_D^{(n)})$ given by

$$\begin{aligned} &- n\log|\varphi'(0)| + D\left(\overline{\mathcal{O}(1)}\cdot\overline{\mathcal{O}(1)}\right) + \left(\sum_{h=1}^{r} b_h \max\{n, u_h D\}\right) \\ &+ n\chi_{[0,\xi]}(n/D)\left(\max_{1\le i\le m} e_i\right) J_\xi(n/D), \end{aligned} \tag{7.7.2}$$

where $\xi \in [0,m]$ is our cutoff parameter from the definition of $\tau^\sharp$ in our estimates, and we set for $s \ge 1$

$$J_\xi(s) := \left(\frac{1}{s}\sum_{j=1}^{\lfloor (s-1)/\max(1,\xi)\rfloor} 1/j\right) + \left(\frac{1}{\lfloor (s + (\xi-1)^+)/\max(1,\xi)\rfloor} - \frac{\xi}{s}\right)^+;$$

and for $s < 1$,

$$J_\xi(s) := \left(1 - \frac{\xi}{s}\right)^+.$$

[31]We do not have such a proof.

By our proof, we only need to show that

$$\sum_{n\in\mathcal{V}_D} \overline{h}(\psi_D^{(n)}) \le \sum_{n=0}^{mD-1} \overline{h}(\psi_D^{(n)}) + o(D^2).$$

To this end, it is sufficient to show that for $n \ge mD$, we have

$$\overline{h}(\psi_D^{(n)}) \le \min_{0\le n'<mD} \overline{h}(\psi_D^{(n')}) + o(D). \tag{7.7.3}$$

We observe $J_\xi(s)$ is a continuous function in s which is piecewise smooth on the intervals of the form $(k\xi+1,(k+1)\xi),((k+1)\xi,(k+1)\xi+1)$, where $k\in\mathbf{N}$. Moreover

$$J'_\xi(s) = -\frac{1}{s^2}\sum_{j=1}^{k} 1/j \le 0$$

on $s\in(k\xi+1,(k+1)\xi)$, and

$$J'_\xi(s) = -\frac{1}{s^2}\left(-\xi+\sum_{j=1}^{k} 1/j\right) \le \xi/s^2$$

on $s\in((k+1)\xi,(k+1)\xi+1)$. Moreover, $J_\xi(s)=0$ for $s\in(0,\min\{1,\xi\}]$. Hence $J_\xi(s)+\xi/s$ is a decreasing function of $s\in\mathbf{R}_{>0}$, and in particular, its $[\xi,\infty)$ maximum is taken at $s=\xi$, with value 1. We analyze the function

$$F(s) := -s\left(\Big(\log|\varphi'(0)|-\sigma_m-(\max_{i=1}^{m} e_i)J_\xi(s)\Big)\right).$$

It is continuous on $s\in\mathbf{R}_{>0}$ and piecewise smooth on the intervals of the form $(k\xi+1,(k+1)\xi),((k+1)\xi,(k+1)\xi+1)$. For $s\ge\xi$, in each of those intervals,

$$\begin{aligned}F'(s) &= -\left(\log|\varphi'(0)|-\sigma_m-\left(\max_{1\le i\le m} e_i\right)J_\xi(s)\right)+\left(\max_{1\le i\le m}(e_i)\right)sJ'(s)\\ &\le -\left(\log|\varphi'(0)|-\sigma_m-\left(\max_{1\le i\le m} e_i\right)J_\xi(s)\right)+\left(\max_{1\le i\le m} e_i\right)\xi/s\\ &\le -\log|\varphi'(0)|+\sigma_m+\left(\max_{1\le i\le m} e_i\right)<0.\end{aligned}$$

Note that since $u_h<m$ and $\xi\le m$, we have for $n\ge mD$:

$$\overline{h}(\psi_D^{(n)}) = -n\log|\varphi'(0)|+D\left(\overline{\mathcal{O}(1)}\cdot\overline{\mathcal{O}(1)}\right)+\sigma_m n+n\left(\max_{1\le i\le m} e_i\right)J_\xi(n/D).$$

Our discussion then shows that $\overline{h}(\psi_D^{(n)})$ is a decreasing function in n on the requisite range $n\ge mD$. In fact, the same argument applies to see that $\overline{h}(\psi_D^{(n)})$ is a decreasing function on the whole $n\in\mathbf{N}$. More precisely, for $n/D\ge\xi$ and $n/D\in[u_h,u_{h+1}]$, we have

$$\begin{aligned}\overline{h}(\psi_D^{(n)}) =& -n\log|\varphi'(0)|+D\left(\overline{\mathcal{O}(1)}\cdot\overline{\mathcal{O}(1)}+\sum_{k=h+1}^{r} u_k b_k\right)+\left(\sum_{k=1}^{h} b_k\right)n\\ &+n\left(\max_{1\le i\le m} e_i\right)J_\xi(n/D).\end{aligned} \tag{7.7.4}$$

Since $\sum_{k=1}^h b_k \leq \sigma_m$, the same argument as above shows that

$$F_h(s) := -s\left(\big(\log|\varphi'(0)| - \sum_{k=1}^{h} b_k - (\max_{i=1}^{m} e_i)J_\xi(s)\big)\right)$$

has negative derivative and hence $\overline{h}(\psi_D^{(n)})$ is a decreasing function in n. For $n/D < \xi$ and $n/D \in [u_h, u_{h+1}]$, we have

$$\overline{h}(\psi_D^{(n)}) = -n\log|\varphi'(0)| + D\left(\overline{\mathcal{O}(1)}\cdot\overline{\mathcal{O}(1)} + \sum_{k=h+1}^{r} u_k b_k\right) + \left(\sum_{k=1}^{h} b_k\right)n.$$

Since $\sum_{k=1}^h b_k \leq \sigma_m < \log|\varphi'(0)|$, we conclude that $\overline{h}(\psi_D^{(n)})$ is a decreasing function.

Thus we obtain (7.7.3), giving a proof of Theorem 7.0.1 free of appeal to the Shidlovsky type theorems from § 3.2, but under the stronger assumption that $\log|\varphi'(0)| > \sigma_m + \max_{i=1}^m(e_i)$. In particular, in a manner free of any of the references in § 3, these remarks already suffice for proving Theorem 2.5.1 except for the clause that $|\varphi'(0)| > e^{\max(\sigma_m, \tau(\mathbf{b}))}$ can be relaxed to $|\varphi'(0)| > e^{\sigma_m}$ when the f_i are *a priori* supposed holonomic.

7.7.5. *Discussion for the* $\mathbf{e} = \mathbf{0}$ *case of Theorem 7.1.6.* As the behavior of the function J_ξ is the main obstacle to devising a clean proof of our general holonomy bounds not relying on functional bad approximability theorems for holonomic functions, and since we do not logically need an alternative proof for any of our applications, we are content (still in the context of explaining how to bypass § 3.2) for Theorem 7.1.6 with demonstrating how to handle the case $\mathbf{e} = \mathbf{0}$ and under the seemingly mild extra condition that

$$m \geq \max_{0\leq k\leq l}\left\{\frac{(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}) - (\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}) + \alpha_k\left(\log|\varphi'_{r_k}(0)| - \sum_{j=1}^{h(k)} b_j\right) - \sum_{j=h(k)+1}^{r} u_j b_j}{\log|\varphi'(0)| - \sigma_m}\right\}, \tag{7.7.6}$$

where, for a given k, we pick the $h(k)$ with $\alpha_k \in [u_{h(k)}, u_{h(k)+1})$, and we recall the convention $\alpha_0 = 0$. In other words, we will show without appealing to § 3.2 that when $\mathbf{e} = \mathbf{0}$ and under the conditions of Theorem 7.1.6, at least one of the dimension bounds (7.1.8) or

$$m \leq \max_{0\leq k\leq l}\left\{\frac{(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}) - (\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}) + \alpha_k\left(\log|\varphi'_{r_k}(0)| - \sum_{j=1}^{h(k)} b_j\right) - \sum_{j=h(k)+1}^{r} u_j b_j}{\log|\varphi'(0)| - \sigma_m}\right\} \tag{7.7.7}$$

is in place. In practice, we have always found that the inequality (7.7.7) is already implied by the contrapositive of the condition inequality (7.1.8), in which case the conclusion (7.1.8) certainly follows.

For $n/D \in [\alpha_k, \alpha_{k+1}]\cap[u_h, u_{h+1}]$ (here h does not need to be $h(k)$ defined above), the proof of Theorem 7.1.6 shows that

$$\overline{h}(\psi_D^{(n)}) = -n\log|\varphi'(0)| - n\log r_k + D(\overline{\mathcal{L}}\cdot\overline{\mathcal{L}}_r) + n\sum_{j=1}^{h} b_j + D\sum_{j=h+1}^{r} u_j b_j, \tag{7.7.8}$$

which, if viewed as a piecewise linear function in $s = n/D$, is continuous on $s \in [0,\infty)$. The local minima can only occur at points of the form $s = \alpha_k$, or over a line segment of slope 0. But for $n/D \geq m$ we use as archimedean height evaluation bound

$$\overline{h}(\psi_D^{(n)}) = -n \log|\varphi'(0)| + D(\overline{\mathcal{L}} \cdot \overline{\mathcal{L}}) + n\sigma_m,$$

which decreases monotonically in $s = n/D$. Now to show (7.7.3), which as in (7.7.1) suffices to bypass the appeal to § 3.2 in our analysis in § 7.4, it is enough to check that

$$\overline{h}(\psi_D^{(mD)}) \leq \min_{0 \leq k \leq l} \overline{h}(\psi_D^{(\alpha_k D)});$$

here as $\alpha_k D$ might not be an integer, by a slight abuse of notation, $\overline{h}(\psi_D^{(\alpha_k D)})$ means replacing n in (7.7.8) by $\alpha_k D$. This unfolds to the definition of the condition (7.7.6).

7.7.9. *Discussion for Theorem 7.1.13.* Under the conditions

$$\begin{aligned} &\xi > u_1 \geq 1, \quad b_1 > \log R \geq 0, \quad \log|\varphi'(0)| > \sigma_m + \max_{1 \leq i \leq m} e_i, \\ &\Gamma(\log R, r, \lambda, \mu) \geq u_1 \log R, \end{aligned} \tag{7.7.10}$$

we show independently of § 3.2 that at least one of the dimension bounds (7.1.14) or

$$m \leq \frac{T(\varphi; r, \lambda, \mu) - \sum_{j=1}^r u_j b_j}{\log|\varphi'(0)| - \sigma_m - (\max_{1 \leq i \leq m} e_i) J_\xi(m)} \tag{7.7.11}$$

is in place. In the practical situations of many applications, usually (7.7.11) is expected to be a smaller bound than (7.1.14). This applies for example to our proof of Theorem A via Example 7.5.9, where the conditions (7.7.10) are met, and the right-hand side of (7.7.11) is negative, whereas the holonomy bound of (7.1.14) is at ~ 13.8527.

Let $\overline{h}_\infty(\psi_D^{(n)})$ and $\overline{h}_{\mathrm{fin}}(\psi_D^{(n)})$ denote the respective main terms of our bounds on $h(\psi_D^{(n)})$, given in (7.5.5), (7.5.4), and (7.5.6) (for the archimedean estimates), and (7.3.13) and (7.3.11) (for the finite estimates), according to the various cases in dependence on n/D. In these notations we have $\overline{h}(\psi_D^{(n)})$ continuous in $s := n/D$, and hence again we only need to ensure (7.7.3).

In the case at hand, our assumptions imply $m \geq \max\{\chi_1, \xi, u_1, \ldots, u_r\}$. For $n \geq mD$, we derive for the left-hand side of (7.7.3):

$$\overline{h}(\psi_D^{(n)}) = -n \left(\log|\varphi'(0)| - \sigma_m - \left(\max_{1 \leq i \leq m} e_i \right) J_\xi(n/D) \right) + DT(r, \lambda, \mu).$$

By the same analysis as in § 7.7.1, we derive that $\overline{h}(\psi_D^{(n)})$ is a decreasing function of n in the range $n \geq \max\{\xi, \chi_1\}D$; therefore for $n \geq mD$, we have

$$\overline{h}(\psi_D^{(n)}) \leq \min_{\max\{\xi,\chi_1\}D \leq n' < mD} \overline{h}(\psi_D^{(n)}).$$

It remains to consider the range $n' < \max\{\xi, \chi_1\}D$. If $\xi > \chi_1$, then for $s' := n'/D \in [\chi_1, \xi] \cap [u_h, u_{h+1}]$, we have

$$\overline{h}(\psi_D^{(n)}) = -n' \left(\log|\varphi'(0)| - \sum_{j=1}^h b_j \right) + D \left(T(r, \lambda, \mu) + \sum_{j=h+1}^r u_j b_j \right),$$

and therefore $\overline{h}(\psi_D^{(n)})$ is a decreasing function in s' in the range $[\chi_1, \xi]$; continuity the gives

$$\overline{h}(\psi_D^{(n)}) \leq \min_{\chi_1 D \leq n' < mD} \overline{h}(\psi_D^{(n)}).$$

Next we take up the range $\xi \le s' < \chi_1$. Here we have for $s' \in [u_h, u_{h+1}]$,

$$\begin{aligned}\overline{h}(\psi_D^{(n')}) =& n'\left(-\log R + \sum_{j=1}^{h} b_j + \left(\max_{1\le i\le m} e_i\right) J_\xi(n/D)\right)\\ &+ D\left(\Gamma(\log R, r, \lambda, \mu) + \sum_{j=h+1}^{r} u_j b_j\right).\end{aligned}$$

Since $u_1 < \xi$, we have $h \ge 1$ for these n' and then $\overline{h}(\psi_D^{(n')})$ is an increasing function of s' due to our assumption $\log R < b_1$ (while, by definition, $J_\xi \ge 0$).

We continue: for $u_1 \le s' < \min\{\xi, \chi_1\}$ (if $u_1 > \chi_1$, then this set is empty and move on to the next case below), we have (the h in the formula may vary depending on s' as above)

$$\overline{h}(\psi_D^{(n')}) = n'(-\log R + \sum_{j=1}^{h} b_j) + D\left(\Gamma(\log R, r, \lambda, \mu) + \sum_{j=h+1}^{r} u_j b_j\right),$$

also an increasing function of s'.

For $\chi_0 \le s' < \min\{u_1, \chi_1\}$,

$$\overline{h}(\psi_D^{(n')}) = -n' \cdot \log R + D\left(\Gamma(\log R, r, \lambda, \mu) + \sum_{j=1}^{r} u_j b_j\right),$$

is a decreasing function of s' due to our assumption $\log R \ge 0$.

Finally, for $0 \le s' < \chi_0$, we have

$$\overline{h}(\psi_D^{(n')}) = D\left(-\lambda(s')^r - \mu s' + \sum_{j=1}^{r} u_j b_j\right).$$

Recall that $\lambda > 0$. If $\mu > 0$, then $\overline{h}(\psi_D^{(n')})$ is a decreasing function. If $\mu \le 0$, the critical point s_0 satisfies that

$$s_0 = \left(\frac{-\mu}{r\lambda}\right)^{1/(r-1)} \le \left(\frac{\log R - \mu}{r\lambda}\right)^{1/(r-1)} = \chi_0,$$

under our assumed conditions; therefore $\overline{h}(\psi_D^{(n')})$ is an increasing function on $[0, s_0]$ and a decreasing function on $[s_0, \chi_0]$.

The above discussion shows that

$$\min_{0\le n'\le mD} \overline{h}(\psi_D^{(n)}) = \min\{\overline{h}(\psi_D^{(0)}), \overline{h}(\psi_D^{(u_1 D)}), \overline{h}(\psi_D^{(mD)})\}.$$

Since

$$\overline{h}(\psi_D^{(0)}) = D\sum_{j=1}^{r} u_j b_j,\ \overline{h}(\psi_D^{(u_1 D)}) = -u_1 D \log R + D\left(\Gamma(\log R, r, \lambda, \mu) + \sum_{j=1}^{r} u_j b_j\right),$$

we have $\overline{h}(\psi_D^{(0)}) \le \overline{h}(\psi_D^{(u_1 D)})$ due to the assumption $\Gamma(\log R, r, \lambda, \mu) \ge u_1 \log R$.

In upshot, the requisite inequality

$$\overline{h}(\psi_D^{(mD)}) \le \min_{0\le n'<mD} \overline{h}(\psi_D^{(n')})$$

boils down to securing that $\overline{h}(\psi_D^{(mD)}) \leq D\sum_{j=1}^r u_j b_j$, which is equivalent to

$$m \geq \frac{T(r,\lambda,\mu) - \sum_{j=1}^r u_j b_j}{\log|\varphi'(0)| - \sigma_m - (\max_{1\leq i\leq m} e_i) J_\xi(m)}.$$

In other words, we have either proved (7.7.3) and hence (7.1.14) holds, or else (7.7.11) holds.

8. The finer holonomy bound with the Bost–Charles integral

This section combines the measure concentration input of § 6 with the Bost–Charles refinements of § 7 to give the most accurate general holonomy bound of all the theorems worked out in our paper. The added strengthening turns out to be zero for the particular applications to Theorems A and C, and the theorem becomes somewhat complicated to state, nevertheless we hope that the principle of the abstract refinement could be useful in future applications of our holonomy bounds.

The theoretical improvement from adding high-dimensional methods to the Bost–Charles calculus is, as far as we were able to tell in the framework of § 2, reflected only in the denominator term $\tau(\mathbf{b};\mathbf{e})$. It is inevitable to ask which of the growth integrals in § 6 versus § 7 is the smaller. In § 8.1, we present a proof by Fedor Nazarov that the Bost–Charles integral is strictly better than the rearrangement integral. This, in particular, implies that Theorem 7.6.4 is more precise than Theorem 6.0.2, granting our heuristic Remark 7.6.7, and at least as far as the stated denominator is concerned in the latter theorem. (Remark 6.6.15 indicates that the multidimensional proof in § 6 can go further than § 7 in the general denominator aspect, and at least as far as our choice of treatment in the present section § 8.) Remark 8.1.17 further down in this section suggests that the difference in the archimedean growth terms is usually very small, implying that little is to be lost from working with the rearrangement integrals. For purposes of sampling and testing the holonomy bounds with different maps φ, the latter integrals have the practical advantage to allow for faster and more reliable numerical computations.

We now proceed to formulating our unifying theorem. The following sums up the sharpest[32] of all the holonomy bounds we prove in this paper.

Theorem 8.0.1. *We relax all assumptions on the denominators in Theorem 6.0.2, and allow for an arbitrary denominators matrix $\mathbf{b} \in M_{m\times r}(\mathbf{R}_{\geq 0})$ with nonnegative coefficients. Define*

$$\begin{aligned} \tau^{\flat\flat}(\mathbf{b}) &:= \limsup_{\epsilon\to 0,\varepsilon\to 0} \limsup_{d\to\infty} \left\{ \frac{1}{d} \limsup_{N\to\infty} \mathrm{Den}_N(\mathbf{b},\epsilon,d,\varepsilon) \right\}, \\ \mathrm{Den}_N(\mathbf{b},\epsilon,d,\varepsilon) &:= \frac{2}{N} \sum_{k\in\mathbf{N}_{>0}} \max_{\mathbf{n}\in P_\varepsilon^d(N),\mathbf{i}\in V_m^d(\epsilon)} \#\{(j,h) \, : \, k \leq b_{i_j,h}\cdot n_j\}, \end{aligned} \tag{8.0.2}$$

where (j,h) runs through $\{1,\ldots,d\}\times\{1,\ldots,r\}$, the set $V_m^d(\epsilon) \subset \{1,\ldots,m\}^d$ is defined as

$$\begin{aligned} V_m^d(\epsilon) := \big\{ \mathbf{i} \in \{1,\ldots,m\}^d \, : \, &\forall i_0 \in \{1,\ldots,m\}, \\ &d/m - \epsilon d < \#\{1\leq j\leq d \mid i_j = i_0\} < d/m + \epsilon d \big\}, \end{aligned}$$

[32] Possibly up to considering other φ in addition to φ_r, as in Theorems 6.0.2 or 7.1.13. Such variations are straightforward to incorporate as well, but we refrain from doing this here.

and $P_{\varepsilon}^{d}(N) \subset [0,N]^d \cap \mathbf{Z}^d$ denotes the subset of those $\mathbf{n}$ for which the normalized $([0,1), \mu_{\text{Lebesgue}})$ discrepancy of $\{n_i/N\}_{i=1}^{d}$ is $\leq \varepsilon$.

Assume either that $\log|\varphi'(0)| > \max\left\{ \sum_{h=1}^{r} \max_{1\leq i\leq m} b_{i,h}, \tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e}) \right\}$*, or that* $\log\varphi'(0) >, \tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e})$ *and all f_i are holonomic.*

Then we have

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\text{Haar}}(z)\mu_{\text{Haar}}(w)}{\log|\varphi'(0)| - (\tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e}))}. \tag{8.0.3}$$

Further, for any subradii sequences $1 = r_l > r_{l-1} > \cdots > r_0 > 0$ as in Theorem 7.1.6 and using the Bost–Charles characteristic $\widehat{T}$ of Definition 7.1.2, we have

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\text{Haar}}(z)\mu_{\text{Haar}}(w) - \frac{1}{m}\sum_{k=1}^{l} \frac{(\widehat{T}(r_k,\varphi)-\widehat{T}(r_{k-1},\varphi))^2}{\log r_k - \log r_{k-1}}}{\log|\varphi'(0)| - (\tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e}))}. \tag{8.0.4}$$

Moreover, if the s_h^* for the given $\{r_k\}$ are defined as in Theorem 7.6.4 then we have

$$m \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\text{Haar}}(z)\mu_{\text{Haar}}(w) - \sum_{h=0}^{l-1} s_h^* \cdot (\widehat{T}(1,\varphi) - \widehat{T}(r_h,\varphi))}{\log|\varphi'(0)| - (\tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e}))}. \tag{8.0.5}$$

Remark 8.0.6. We observe that $\tau^{\flat\flat}(\mathbf{b}) \in [0,\infty)$ by definition. More precisely, we have the trivial bound $\tau^{\flat\flat}(\mathbf{b}) \leq \sum_{h=1}^{r} \max_{i=1}^{m}\{b_{i,h}\}$.

For the denominator matrices $\mathbf{b}$ of the form considered throughout §§ 6–7, we have $\tau^{\flat\flat}(\mathbf{b}) = \tau^{\flat}(\mathbf{b})$. Indeed, writing t for the continuous limit of the discrete variable k/N, we have in the setup of Theorem 6.0.2 $\tau^{\flat\flat}(\mathbf{b})$ bounded above by

$$\sum_{h=1}^{r} \limsup_{\epsilon,\varepsilon\to 0} \limsup_{d\to\infty} \left\{ \frac{1}{d} \limsup_{N\to\infty} \left\{ \frac{2}{N} \sum_{k\in\mathbf{N}_{>0}} \max_{\mathbf{n}\in P_{\varepsilon}^{d}(N), \mathbf{i}\in V_m^d(\epsilon)} \#\{(j,h) \,:\, k \leq b_{i_j,h}\cdot n_j\} \right\} \right\}$$
$$\leq 2\sum_{h=1}^{r} \int_0^{b_h} \min\{1-u_h/m, 1-t/b_h\}\, dt = \sum_{h=1}^{r} \left(b_h - \frac{b_h u_h^2}{m^2} \right) = \tau^{\flat}(\mathbf{b}),$$

where the inequality stems from the observation that the restriction to the *balanced* $\mathbf{i}$ (meaning: each $i_0 \in \{1,\ldots,m\}$ occurs with the same asymptotic frequency $1/m$) supplies the upper bound constraint $1 - u_h/m$ in the integrand on the second line, while the restriction to the *balanced* $\mathbf{n}$ (meaning: the components set $\{n_j\}$ takes asymptotically the uniform distribution on $[0,N]$) supplies the upper bound constraint $1 - t/b_h$ in that integrand. Moreover, equalities are reached in the case where both $\mathbf{n}$ and $\mathbf{i}$ are arranged in non-decreasing order: $n_1 \leq \cdots \leq n_d$ and $i_1 \leq \cdots \leq i_d$. This proves that $\tau^{\flat\flat}(\mathbf{b}) = \tau^{\flat}(\mathbf{b})$ under the standing assumptions throughout § 6 and § 7. (See also Lemma 6.6.7 and Remark 6.6.14.) △

Example 8.0.7. Here is a simple example to illustrate that, for a given $\mathbf{b}$, if one lets $\mathbf{b}'$ range over all arrays which dominate $\mathbf{b}' \geq \mathbf{b}$ coefficient-wise and which additionally meet the constraints of Theorem 6.0.2, the inequality in $\tau^{\flat\flat}(\mathbf{b}) \leq \tau^{\flat}(\mathbf{b}')$ can be strict.

Consider $\mathbf{b} = [0,1,2]^{\mathrm{t}}$. In order to use $\tau^{\flat}$, the optimal choice of $\mathbf{b}'$ is to take $\mathbf{b}' = [0,2,2]^{\mathrm{t}}$ and we have $\tau^{\flat}(\mathbf{b}') = \frac{16}{9}$. On the other hand, we have $\tau^{\flat\flat}(\mathbf{b}) = \frac{5}{3} = \frac{15}{9}$.

Indeed, writing $t := k/N$, we have:

$$\tau^{\flat\flat}(\mathbf{b}) = 2\left(\int_0^{2/3} \frac{2}{3}\,dt + \int_{2/3}^{1}\left(1 - t + \frac{1}{3}\right)\,dt + \int_1^{4/3} \frac{1}{3}\,dt + \int_{4/3}^{2}\left(1-\frac{t}{2}\right)\,dt\right).$$

This is to be compared with

$$\tau^{\flat}(\mathbf{b}') = 2\int_0^2 \min\left\{\frac{2}{3}, 1 - \frac{t}{2}\right\}\,dt.$$

To explain the difference between two formulas, we notice that in the range $t \in [2/3, 4/3]$, for every $\mathbf{n}$ to be considered in the definition of $\tau^{\flat\flat}(\mathbf{b})$, there are at most $(\max\{0, 1-t\} + o(1))d$ among the n_j with $i_j = 2$ (corresponding to $b_{2,1} = 1$) to contribute to $\{(j,1) \mid k \leq n\}$, and there are at most $(1/3 + o(1))d$ among the n_j with $i_j = 3$ (corresponding to $b_{3,1} = 2$) to contribute to $\{(j,1) \mid k \leq 2n\}$ and both bounds can be reached with suitable choice of $\mathbf{n}$. △

8.1. **Comparison of the Bost–Charles and the Rearrangement integrals.** This section, due entirely to Fedor Nazarov, treats the clean comparison of the two integrals — beneath the empirical observation that they are practically the same in the situations we encounter in § A, as well as in practice for most of the multivalent cases. This theorem strictly speaking is not used for any of our proofs in the paper, and hence it can be omitted on a first (and on a second) reading.

Basic Remark 8.1.1. As the considerations that follow rely on the potential theory in the plane, consider first the easier situation on the simplest of all the Lie groups: the circle $\mathbf{T}$. The integrable function $G(z) := \log\frac{1}{|1-z|}$, $G : \mathbf{T} \to \mathbf{R} \cup \{\infty\}$ has the nonnegative Fourier coefficients

$$\widehat{G}(n) := \int_{\mathbf{T}} G(z) z^{-n}\,\mu_{\mathrm{Haar}}(z) = \int_{\mathbf{T}} z^{-n} \log\frac{1}{|1-z|}\,\mu_{\mathrm{Haar}}(z) = \begin{cases} 0, & \text{if } n = 0; \\ \dfrac{1}{2|n|}, & \text{if } n \neq 0. \end{cases} \tag{8.1.2}$$

By the general Bochner theorem, the positivity of the Fourier transform implies that $G(z)$ is a positive-definite function on the locally compact abelian group $\mathbf{T}$: that is, $\iint_{\mathbf{T}^2} G(zw^{-1})\,\nu(z)\nu(w) \geq 0$ for all *reasonable* signed measures ν on $\mathbf{T}$. *Reasonable* here may be taken to mean $\nu = \nu^+ - \nu^-$ with finite positive measures $\nu^\pm$ satisfying $\int_{\mathbf{T}} G\,\nu^\pm < \infty$. For any such signed measure ν, this computation shows more precisely that

$$\begin{aligned} I(\nu) &:= \iint_{\mathbf{T}^2} \log\frac{1}{|z-w|}\,\nu(z)\nu(w) \\ &= \iint_{\mathbf{T}^2} \log\frac{1}{|1-zw^{-1}|}\,\nu(z)\nu(w) = \sum_{n \in \mathbf{Z}\smallsetminus\{0\}} \frac{|\widehat{\nu}(n)|^2}{2|n|} \geq 0, \end{aligned} \tag{8.1.3}$$

where manifestly the equality holds if and only if the Fourier transform $\widehat{\nu}$ is a scalar multiple of the Dirac mass at $0 \in \widehat{\mathbf{T}} = \mathbf{Z}$, and that in turn is the case if and only if ν is a scalar multiple of the Haar measure μ_{Haar}. This reflects the basic potential theory on the circle.

On the circle $|z| = R$ of arbitrary radius, the left-hand side of (8.1.3) scales by the additive summand $(\log R)\left(\int \nu\right)^2$, and so whereas for $R > 1$ the inequality (8.1.3) is false for arbitrary measures ν supported by that circle, it continues to be in place for the measures that are *balanced* in the sense that $\int \nu = \nu(\mathbf{C}) = 0$. As the logarithmic kernel is also invariant under additive translations, the latter remark continues to hold for balanced measures carried by any circle in $\mathbf{C}$. △

As we review next, the positivity of the energy integral is a completely general fact about balanced measures on $\mathbf{C}$.

8.1.4. *The energy principle.* In the Newtonian gravitational field created by a point mass at the origin $\mathbf{0} \in \mathbf{R}^n$, the potential energy function $U : \mathbf{R}^n \smallsetminus \{\mathbf{0}\} \to \mathbf{R}$ is determined by the distributional Laplace equation

$$\Delta U := \sum_{i=1}^{n} \frac{\partial^2 U}{\partial x_i^2} = -\delta_{\mathbf{0}},$$

namely as the fundamental solution $U := U_n^2$ of that equation, where more generally it is useful to consider the *Riesz potential* [Rie38] defined by

$$U_n^\alpha(\mathbf{x}) := \begin{cases} \dfrac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{n}{2}}} \log \dfrac{1}{\|\mathbf{x}\|}, & \text{for } n = \alpha; \\ \dfrac{\Gamma\left(\frac{n-\alpha}{2}\right)}{2^\alpha \Gamma\left(\frac{\alpha}{2}\right) \pi^{\frac{n}{2}}} \dfrac{1}{\|\mathbf{x}\|^{n-\alpha}}, & \text{for } n > \alpha > 0. \end{cases} \tag{8.1.5}$$

Here, $\|\mathbf{x}\| := \sqrt{\sum_{i=1}^n x_i^2}$ is the Euclidean distance function in $\mathbf{R}^n$. The gravitational potential $U = U_n^2$ is rotationally invariant and defines a kernel function $k(\mathbf{x},\mathbf{y}) := U_n^2(\mathbf{x}-\mathbf{y})$, which by definition is furthermore translationally invariant. For $n > 2$ this kernel is also positive-definite, by the fundamental formula of Frostman and Marcel Riesz [Rie38, § I.3], which generalizes the Dirichlet energy integral $\frac{1}{2}\iint \|\nabla F\|^2 \, \mathrm{dvol}$, and is tantamount to the computation [Den50, Sch66, Lan72, NS91] of the distributional Fourier transform of U_n^2:

$$\begin{aligned} I(\nu) &:= \iint_{\mathbf{R}^n \times \mathbf{R}^n} k(\mathbf{x},\mathbf{y}) \, \nu(\mathbf{x}) \nu(\mathbf{y}) = \iint_{\mathbf{R}^n \times \mathbf{R}^n} U_n^2(\mathbf{x}-\mathbf{y}) \, \nu(\mathbf{x}) \nu(\mathbf{y}) \\ &= \int_{\mathbf{R}^n} \left(U_n^1 * \nu\right)^2 \, \mu_{\mathrm{Lebesgue}}. \end{aligned} \tag{8.1.6}$$

This is the spatial analogue of the energy formula (8.1.3) for the circle. To be more precise, this formula gives the strict positivity $I(\nu) > 0$ of the energy of any nonzero compactly supported signed measure $\nu = \nu^+ - \nu^-$ on $\mathbf{R}^n$ expressible as the difference of two finite positive Borel measures ν^+, ν^- of finite energies $I(\nu^+), I(\nu^-) < \infty$. For the logarithmic kernel (the case $n = 2$, whose proper physical interpretation is rather in electrostatics on a plate), Riesz observed [Rie38, § I.4] that the analogy becomes almost perfect upon additionally requiring the signed measure ν to be *balanced*: $\nu(\mathbf{R}^2) = 0$. The *energy principle* states that for balanced signed measures subject to the above regularity conditions (with the balancing condition being only required in the case $n = 2$), the energy $I(\nu) \geq 0$ is nonnegative, and equality holds if and only if $\nu = 0$. This refines the uniqueness theorem for the equilibrium probability measure of a compact. The $n = 2$ case, which is the one of relevance to us, is treated in detail in [Hil62, Theorem 16.4.2].

8.1.7. *Fuglede's inequality.* For completeness, we summarize the more general situation due to Fuglede [Fug60] for non-balanced measures. This is a different generalization of the $n = 2$ case, this time to the logarithmic kernel

$$k(\mathbf{x},\mathbf{y}) := U_n^n(\mathbf{x}-\mathbf{y}) = \frac{\Gamma(n/2)}{2\pi^{n/2}} \log \frac{1}{\|\mathbf{x}-\mathbf{y}\|}$$

on $\mathbf{R}^n$. For this kernel, the analog of the energy formula (8.1.6) for the case of the balanced measures is also due to Riesz [Rie38, § I.4]:

$$\int_{\mathbf{R}^n} \nu = 0 \quad \Longrightarrow \quad I(\nu) = \int_{\mathbf{R}^n} \left(U_n^{n/2} * \nu\right)^2 \mu_{\text{Lebesgue}} \geq 0,$$

and more generally, Fuglede [Fug60, § 4] proves the sharp energy lower bound

$$I(\nu) \geq \log\left(\frac{a_n}{R}\right) \cdot \left(\int_{\mathbf{R}^n} \nu\right)^2 \tag{8.1.8}$$

for all signed measures ν on $\mathbf{R}^n$ expressible as $\nu = \nu^+ - \nu^-$ with $\nu^\pm$ finite positive measures of convergent energy integrals $I(\nu^\pm) < \infty$ and having $\operatorname{supp}(\nu) \subseteq \{\|\mathbf{x}\| \leq R\}$, and with the optimal constant a_n being precisely

$$a_n := \begin{cases} \exp\left(\dfrac{1}{2} + \ldots + \dfrac{1}{n-4} + \dfrac{1}{n-2}\right), & \text{for } n \text{ even;} \\ \exp\left(\dfrac{1}{1} + \ldots + \dfrac{1}{n-4} + \dfrac{1}{n-2} - \log 2\right), & \text{for } n \text{ odd.} \end{cases}$$

The cases $a_1 = 1/2$ and $a_2 = 0$ (for the case $n = 2$ of relevance to us) are already in de la Vallee-Poussin [dlVP49, § 47]. In any case, (8.1.8) certainly implies the requisite positivity $I(\nu) \geq 0$ for all (reasonable) *balanced* measures, and more generally, for all measures supported by a sufficiently small ball.

8.1.9. *Michelli's criterion for positive-definite kernels.* A different generalization, for which we refer to [Mat97] and the references there, admits an arbitrary kernel of the form $k(\mathbf{x}, \mathbf{y}) = U(\|\mathbf{x} - \mathbf{y}\|)$, where $U(t) \in C^\infty(\mathbf{R}_{>0})$ obeys $(-1)^n \left(\frac{d}{dt}\right)^n U(t) \geq 0$ for all $n \geq n_0$, and now the additional ("balancing") constraints on the compactly supported signed measure $\nu = \nu^+ - \nu^-$ with $I(\nu^\pm) < \infty$ on $\mathbf{R}^d$ being $\int_{\mathbf{R}^d} \mathbf{x}^{\mathbf{m}}\, \nu(\mathbf{x}) = 0$ for all $\mathbf{m} \in \mathbf{N}^d$ with $|\mathbf{m}| < n_0$. The condition on $U(t)$ is equivalent to the existence of an integral representation for $\left(\frac{d}{dt}\right)^{n_0} U(t) = \int_0^\infty e^{-tu}\, \mathrm{d}\alpha(t)$ as a Laplace–Stieltjes transform of a *positive* Borel measure $\mathrm{d}\alpha$ on $\mathbf{R}_{>0}$.

8.1.10. *Intersection pairing and signature.* Another way of informally summarizing this discussion[33] is to say that the infinite-dimensional quadratic form

$$\langle \mu, \nu \rangle := \iint_{\mathbf{R}^n \times \mathbf{R}^n} k(\mathbf{x}, \mathbf{y})\, \mu(\mathbf{x}) \nu(\mathbf{y})$$

has one '$-$' sign on the space of *reasonable* (non-balanced) signed measures on $\mathbf{R}^n$. For the case $n = 2$ of relevance to the rest of § 8.1, it could be interesting to know if a more precise connection could be drawn to the arithmetic Hodge index formula in Arakelov theory and the computations in [BC22, § 5] that led to the Bost–Charles double integral. Is there a proof of Nazarov's inequality (Proposition 8.1.13 below) that works directly into the arithmetic intersection theory framework of [BC22]? A basic remark in the algebraic model is that for any two line bundles L, M on a polarized normal projective algebraic surface, if $(L.L) = (M.M)$ and $\deg L = \deg M$, then $(L.L) \leq (L.M)$ following from the Hodge index theorem for the line bundle $L \otimes M^{-1}$.

---

[33] This is taking $n_0 := 1$ in § 8.1.9 if we are to include the more general kernels $k(x, y)$ there. For arithmetic geometry, the case of relevance is $n = 2$ and the kernel $k(z, w) = -\log|z - w|$.

8.1.11. *Nazarov's inequality.* Consider now the case $n=2$ as the complex plane $\mathbf{R}^2 \cong \mathbf{C}$. Then the rotation group is realized by the unitary transformations $t_a : z \mapsto az$, indexed by the circle points $a \in \mathbf{T}$, and our rotationally-invariant, positive-definite kernel is given by $k(z,w) = \frac{1}{2\pi} \log \frac{1}{|z-w|}$ on the $\mathbf{C}$-linear space of balanced signed measures on $\mathbf{C}$ with the regularity conditions we described. We consider a non-constant continuous function $\varphi : \mathbf{T} \to \mathbf{C}$, and for any $a \in \mathbf{T}$ we apply the energy principle $I(\nu) \geq 0$ to the balanced signed measure $\nu = \nu_a := \varphi_*(\mu_{\mathrm{Haar}}) - t_a^* \varphi_*(\mu_{\mathrm{Haar}})$. As the kernel $k(z,w)$ is symmetric and rotationally invariant, the resulting inequality rewrites as

$$\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) \leq \iint_{\mathbf{T}^2} \log|a\varphi(z)-\varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w), \tag{8.1.12}$$

for any $a \in \mathbf{T}$, and with equality holding if and only if $\nu_a = 0$. Integrating over $a \in \mathbf{T}$ we get:

Proposition 8.1.13 (Nazarov). *For any continuous function $\varphi : \mathbf{T} \to \mathbf{C}$,*

$$\begin{aligned}\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) &\leq \iint_{\mathbf{T}^2} \log\max\left(|\varphi(z)|, |\varphi(w)|\right)\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) \\ &= \int_0^1 2t \cdot \left(\log|\varphi(e^{2\pi i t})|\right)^*\, dt,\end{aligned}$$

where g^* denotes the increasing rearrangement (2.4.1) *of a continuous function $(0,1) \to \mathbf{R}$. Furthermore, equality holds if and only if $\varphi(z) = cz^m$ for some $c \in \mathbf{C}$ and $m \in \mathbf{Z}$.*

Proof. Using the Poisson formula

$$\int_{\mathbf{T}} \log|ax-y|\, \mu_{\mathrm{Haar}}(a) = \log\max\left(|x|,|y|\right)$$

in the termwise integration $\int_{\mathbf{T}} I(\nu_a)\, \mu_{\mathrm{Haar}}(a) \geq 0$ of (8.1.12). The equality requires $\nu_a = 0$ for almost all $a \in \mathbf{T}$, hence that $\varphi_*(\mu_{\mathrm{Haar}})$ is rotationally-invariant, and hence that the supporting loop $\varphi(\mathbf{T}) \subset \mathbf{C}$ is rotationally-invariant and therefore a centered circle, and that $\varphi^*(\mu_{\mathrm{Haar}})$ is a scalar multiple of the Haar measure of that circle. ☐

In particular, we get a clean proof that the Bost–Charles integral is always strictly majorized by the doubled Nevanlinna characteristic that we have in [CDT24] (and not merely by the slightly larger doubled Ahlfors–Shimizu characteristic

$$\int_{\mathbf{T}} \log\sqrt{1+|\varphi|^2}\, \mu_{\mathrm{Haar}}$$

noted in [BC22, Prop. 5.4.5], which is a more basic estimate following simply by the trivial pointwise inequality $|x-y|^2 \leq (1+|x|^2)(1+|y|^2)$):

Corollary 8.1.14. *Every continuous function $\varphi : \mathbf{T} \to \mathbf{C}$ satisfies*

$$\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) \leq 2\int_{\mathbf{T}} \log^+|\varphi|\, \mu_{\mathrm{Haar}}.$$

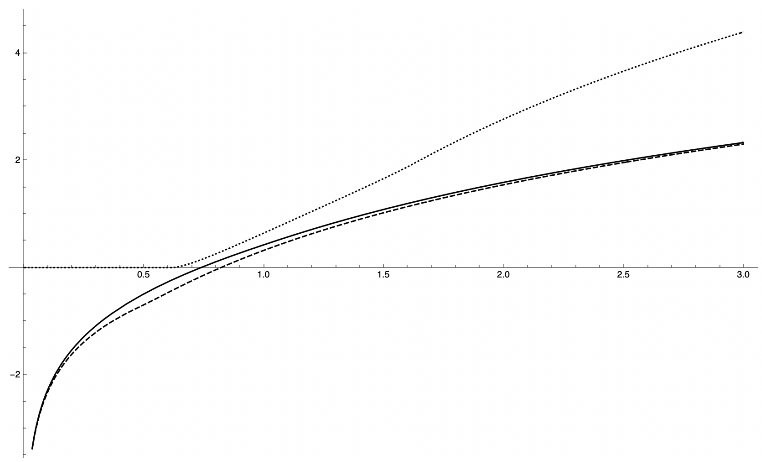

FIGURE 8.1.15. The plots of the Rearrangement (in solid) and the Bost–Charles integrals (dashed) for the bivalent function $\varphi_r(z) = (rz) - (rz)^2$. Above them (dotted), the plot of $2\int_{\mathbf{T}} \log^+ |\varphi_r|\, \mu_{\mathrm{Haar}}$.

Example 8.1.16. Consider the function $\varphi_r(z) := rz - (rz)^2$ with the varying radius r. For $r \geq 1$, the Bost–Charles integral amounts to

$$\iint_{\mathbf{T}^2} \log |\varphi_r(z) - \varphi_r(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) = 2\log r + \int_{\mathbf{T}} \log^+ \left|\frac{r}{1-rz}\right| \mu_{\mathrm{Haar}}(z)$$
$$= 2\log r + \frac{1}{\pi r} + \frac{1}{72\pi r^2} + \dots.$$

In comparison, a computation reveals the rearrangement integral as the explicit function

$$\int_0^1 2t \cdot (\log |\varphi_r(e^{2\pi i t})|)^*\, dt = 2\log r + \frac{1}{2\pi^2}\left(8\,\mathrm{Li}_3\left(1/r\right) - \mathrm{Li}_3\left(1/r^2\right)\right)$$
$$= 2\log r + \frac{4}{\pi^2 r} + \frac{4}{27\pi r^2} + \dots.$$

The comparison is pretty tight for most values of r, as illustrated by Figure 8.1.15. In contrast, the Nevanlinna characteristic upper bound by

$$2\int_{\mathbf{T}} \log^+ |\varphi_r|\, \mu_{\mathrm{Haar}} = 2\int_{|z|=r} \log^+ |z - z^2|\, \mu_{\mathrm{Haar}}(z)$$
$$= 4\log r \text{ when } r \geq \frac{\sqrt{5}+1}{2}$$

is quite crude. △

Remark 8.1.17. The asymptotic equivalence of the two growth characteristic integrals at a "big" radius $|z| = r$ (as observed in Example 8.1.16 and Figure 8.1.15) seems to be a fairly general feature that reflects the near-rotational invariance of the φ_* pushforward of the uniform measure μ_{Haar} of the expanding circle $|z| = r$, considering the proof of Proposition 8.1.13 via the slice-by-slice inequality (8.1.12). An example "in action" is in our proof of Theorem C in § 14.5. In this situation, the

Bost–Charles integral (A.6.1) compares with the rearrangement integral (A.6.3) as follows:

$$9.963 \sim \iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z,w) \leq \int_0^1 2t\cdot(\log|\varphi(e^{2\pi it})|)^*\,dt \sim 9.972,$$

an improvement on the order of merely a tenth of a percent. (In contrast, the Nevanlinna characteristic quantity $2T(\varphi) = 2\int_{\mathbf{T}} \log^+|\varphi|\,\mu_{\mathrm{Haar}}$ from (B.0.1) is in this case as big as ~ 14.08.)

Another example of this kind can be seen by comparing the upper bounds occurring in the two proofs of Theorem A coming from Theorem 6.0.2 and Theorem 7.6.4 respectively. The proof via Theorem 6.0.2 (with two division points, see § 13) gives a bound in terms of rearrangement integrals of the form $m < 13.731$. On the other hand, Theorem 7.6.4 gives a bound in terms of arithmetic intersection numbers which are expressible as integrals which are slight generalizations of the Bost–Charles integrals (see Lemma 7.4.5). With an analogous choice of division points, this leads to the bound $m < 13.679$ (see Example 7.6.8, in particular Equation 7.6.9), an improvement of less than half a percent. This comparison is not literally an example of Proposition 8.1.13 because of the slightly modified forms of both the integrals and the bounds arising from the convexity argument. However, it accurately reflects the amount of improvement between the rearrangement integrals and Bost–Charles integrals that we observed numerically without convexity. △

8.2. **Proof of Theorem 8.0.1.** We firstly prove (8.0.4), which contains (8.0.3) as a special case. We will discuss at the end how to modify the proof to get (8.0.5).

For the following, we fix an $\epsilon > 0$ and the number of variables d. Only at the end of the proof we will let, firstly, $d \to \infty$, followed by $\epsilon \to 0$. To compare to § 6 and (6.0.10), we remark that the bounds in Theorem 8.0.1 only have $\tau^{\flat\flat}(\mathbf{b})$ in the denominator using a high dimensional equidistribution feature, while the other terms are the same as those in the one dimensional bounds in § 7. Therefore, to prove Theorem 8.0.1, we only need to incorporate the feature of a balanced index $\mathbf{i} = (i_1,\dots,i_d)$ in $\prod_{j=1}^d f_{i_j}(x_j)$, while we do not need an equidistributed degree $\mathbf{k}$ in $\mathbf{x}^{\mathbf{k}}$ as in § 6.2. This motivates the following choice for the Euclidean lattice to underlie our auxiliary evaluation module.

8.2.1. *The Euclidean lattice.* Consider the free $\mathbf{Z}$-module

$$E_D := \bigoplus_{\mathbf{i}\in V_m^d(\epsilon)} f_{\mathbf{i}}(\mathbf{x})\,\mathbf{Z}[1/x_1,\dots,1/x_d]_{\leq D}, \tag{8.2.2}$$

where we recall that $V_m^d(\epsilon) \subset \{1,\dots,m\}^d$ is defined (depending on the fixed small positive constant ϵ) as

$$\begin{aligned} V_m^d(\epsilon) = \big\{\mathbf{i}\in\{1,\dots,m\}^d \,:\, &\forall i_0\in\{1,\dots,m\},\\ &d/m-\epsilon d < \#\{1\leq j\leq d \mid i_j = i_0\} < d/m+\epsilon d\big\}; \end{aligned}$$

and, as usual, $f_{\mathbf{i}}(\mathbf{x}) := \prod_{j=1}^d f_{i_j}(x_j)$. Here, $\mathbf{Z}[1/x_1,\dots,1/x_d]_{\leq D}$ denotes the free $\mathbf{Z}$-module consisting of integer polynomials in $1/x_1,\dots,1/x_d$, all of whose partial degrees — with respect to each $1/x_j$ — are at most D.

By construction, $\operatorname{rank} E_D = (D+1)^d \cdot \# V_m^d(\epsilon)$. By Theorem 4.2.1 (actually, the weak law of large numbers suffices here), we have

$$\lim_{d\to\infty} \frac{\# V_m^d(\epsilon)}{m^d} = 1, \text{ thus } \lim_{d\to\infty}\lim_{D\to\infty} \frac{\operatorname{rank} E_D}{m^d D^d} = 1.$$

In order to endow E_D with the suitable norm, we consider the smooth projective arithmetic scheme $\mathcal{X} := (\mathbf{P}^1_{\mathbf{Z}})^d$ and the natural very ample line bundle $\mathcal{L} := \otimes_{j=1}^d \operatorname{pr}_j^* \mathcal{O}(1)$ on $\mathcal{X}$, where $\operatorname{pr}_j : \mathcal{X} \to \mathbf{P}^1_{\mathbf{Z}}$ denotes the projection onto the j-th component. Then we can identify $\Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})$ with $\mathbf{Z}[1/x_1, \ldots, 1/x_d]_{\le D}$, where $x_j := Y_j/Z_j$ is an affine coordinate of the j-th $\mathbf{P}^1_{\mathbf{Z}} = \operatorname{Proj} \mathbf{Z}[Y_j, Z_j]$.

Recall that we are given a holomorphic map $\varphi : (\overline{\mathbf{D}}, 0) \to (\mathbf{P}^1(\mathbf{C}), 0)$, where "0" can mean $x_j = 0$ for each copy of $\mathbf{P}^1_{\mathbf{C}}$ in $\mathcal{X}_{\mathbf{C}}$. The Bost–Charles metric from § 7.3.1 is thus defined using φ on every factor $\operatorname{pr}_j^* \mathcal{O}(1)$. This induces a Hermitian line bundle structure $\overline{\mathcal{L}} = (\mathcal{L}, \|\cdot\|_{\overline{\mathcal{L}}})$ on $\mathcal{L}$, and in turn, as in § 7.3.3, a Euclidean lattice $\Gamma_{L^2}\left(\mathcal{X}, \nu; \overline{\mathcal{L}}^{\otimes D}\right)$ after we fix a smooth probability measure ν on $(\mathbf{P}^1)^d(\mathbf{C})$. The choice of ν is immaterial to the proof, since $D \to \infty$ for the fixed ν and we only need to study the asymptotic leading order term given by the arithmetic Hilbert–Samuel formula for $\widehat{\deg}\, \Gamma_{L^2}\left(\mathcal{X}, \nu; \overline{\mathcal{L}}^{\otimes D}\right)$. For concreteness, we pick ν to be the smooth measure

$$\nu := \bigwedge_{j=1}^d \operatorname{pr}_j^* \omega_{\mathrm{FS}} = \left(\frac{\sqrt{-1}}{2\pi}\right)^d \frac{dz_1 \wedge \cdots \wedge dz_d \wedge d\bar{z}_1 \wedge \cdots \wedge d\bar{z}_d}{\prod_{j=1}^d (1+|z_j|^2)^2}$$

with $\omega_{\mathrm{FS}} = \frac{\sqrt{-1}}{2\pi} \frac{dz \wedge d\bar{z}}{(1+|z|^2)^2}$ being the Fubini–Study form on $\mathbf{P}^1(\mathbf{C})$. Then, as in § 7.3.3, the Euclidean norm $\|\cdot\|$ on $\Gamma\left(\mathcal{X}, \overline{\mathcal{L}}^{\otimes D}\right)$ is defined by

$$\|s\| := \sqrt{\int_{\mathcal{X}(\mathbf{C})} \|s\|^2_{\overline{\mathcal{L}}}\, \nu}. \tag{8.2.3}$$

Just as in § 7.3.3, we take the orthogonal direct sum (8.2.2) of the Euclidean lattices

$$\mathbf{Z}[1/x_1, \ldots, 1/x_d]_{\le D} \cong \Gamma(\mathcal{X}, \mathcal{L}^{\otimes D}) \subset \Gamma\left(\mathcal{X}, \mathcal{L}^{\otimes D}\right)_{\mathbf{R}}$$

induced from the above norm (8.2.3) on $\Gamma\left(\mathcal{X}, \overline{\mathcal{L}}^{\otimes D}\right)$. We use $\overline{E}_D = (E_D, \|\cdot\|)$ to denote this Euclidean lattice.

8.2.4. *Arithmetic degree.* The asymptotic calculation of the arithmetic degrees of direct images to $\operatorname{Spec} \mathbf{Z}$ is the subject of the arithmetic Hilbert–Samuel formula:

Lemma 8.2.5. *As $D \to \infty$, we have the following asymptotics of arithmetic degrees:*

$$\begin{aligned} \widehat{\deg}\, \Gamma_{L^2}\left(\mathcal{X}, \nu; \overline{\mathcal{L}}^{\otimes D}\right) &= \frac{d}{2}\left(\overline{\mathcal{O}(1)} \cdot \overline{\mathcal{O}(1)}\right) D^{d+1} + o(D^{d+1}), \\ \widehat{\deg}\, \overline{E}_D &= \frac{d \# V_m^d(\epsilon)}{2}\left(\overline{\mathcal{O}(1)} \cdot \overline{\mathcal{O}(1)}\right) D^{d+1} + o(D^{d+1}). \end{aligned} \tag{8.2.6}$$

Proof. As Euclidean lattices, $\Gamma(\mathcal{X}, \overline{\mathcal{L}}^{\otimes D}) \cong \otimes_{j=1}^d \Gamma\left(\mathbf{P}^1_{\mathbf{Z}}, \overline{\mathcal{O}(1)}^{\otimes D}\right)$, where the norm on each factor $\Gamma\left(\mathbf{P}^1_{\mathbf{Z}}, \overline{\mathcal{O}(1)}^{\otimes D}\right)$ is the L^2-norm using the Hermitian line bundle

$\overline{\mathcal{O}(1)}$ and the Fubini–Study form ω_{FS} on $\mathbf{P}^1(\mathbf{C})$. Therefore, by Lemma 7.2.7, we have

$$\widehat{\deg}\,\Gamma\left(\mathcal{X},\overline{\mathcal{L}}^{\otimes D}\right)=d(D+1)^{d-1}\widehat{\deg}\,\Gamma\left(\mathbf{P}^1_{\mathbf{Z}},\overline{\mathcal{O}(1)}^{\otimes D}\right).$$

Now the first assertion follows from the arithmetic Hilbert–Samuel formula (7.3.4). The second assertion follows from the first one by (7.2.4)). □

Remark 8.2.7. A proof of this calculation is also a consequence of the general arithmetic Hilbert–Samuel formula in Krull dimension $d+1$. In [Zha95, Theorem 1.4], the Hilbert–Samuel formula is proven for an arithmetic variety of any dimension and an ample Hermitian line bundle with smooth metric of pointwise non-negative Chern form. Using the idea in [Bos99, §5] and [BC22, §§3–4], the same formula continues to hold for an ample line bundle with a $\mathcal{C}^{\mathrm{b}\Delta}$ Hermitian metric of pointwise non-negative Chern form over an arithmetic variety of any dimension. Thus we may also deduce Lemma 8.2.5 directly from the arithmetic Hilbert–Samuel formula for $(\mathcal{X},\overline{\mathcal{L}})$:

$$\widehat{\deg}\,\Gamma_{L^2}\left(\mathcal{X},\nu;\overline{\mathcal{L}}^{\otimes D}\right)=\frac{\overline{\mathcal{L}}^{d+1}}{(d+1)!}D^{d+1}+o(D^{d+1}).$$

Here, $\overline{\mathcal{L}}^{d+1}$ denotes the arithmetic self-intersection number $\widehat{c}_1(\overline{\mathcal{L}})^{d+1}.[\mathcal{X}]$. In our situation, we write $\overline{\mathcal{L}}=\otimes_{j=1}^d\mathrm{pr}_j^*\,\overline{\mathcal{O}(1)}$, and expand the self-intersection number by multilinearity. The only nonzero terms come from crossing all but one of the $\mathrm{pr}_j^*\,\overline{\mathcal{O}(1)}$ factors once and the remaining $\mathrm{pr}_j^*\,\overline{\mathcal{O}(1)}$ factor twice. By the projection formula, that gives us

$$\overline{\mathcal{L}}^{d+1}=d\binom{d+1}{2}(d-1)!\left(\overline{\mathcal{O}(1)}\cdot\overline{\mathcal{O}(1)}\right),$$

and we recover Lemma 8.2.5 by this perspective also. △

8.2.8. *Evaluation filtration.* Writing $X:=\mathcal{X}_{\mathbf{Q}}$, we can identify $\operatorname{Spf}\mathbf{Q}[\![\mathbf{x}]\!]=\widehat{X}_{\mathbf{0}}$, giving in particular elements

$$f_{\mathbf{i}}=f_{\mathbf{i}}(\mathbf{x})\in\Gamma(\widehat{X}_{\mathbf{0}},\mathcal{O}_{\widehat{X}_{\mathbf{0}}}).$$

The space $\Gamma\left(\widehat{X}_{\mathbf{0}},\mathcal{L}^{\otimes D}\right)$ of global sections of $\mathcal{L}^{\otimes D}|_{\widehat{X}_{\mathbf{0}}}$ is then naturally identified with

$$\Gamma\left(\widehat{X}_{\mathbf{0}},\mathcal{L}^{\otimes D}\right)=\mathbf{x}^{-D}\mathbf{Q}[\![\mathbf{x}]\!]=:F_{\mathbf{Q}},$$

where $\mathbf{x}^{-D}\mathbf{Q}[\![\mathbf{x}]\!]$ denotes the $\mathbf{Q}$-vector space generated by the $\mathbf{x}^{\mathbf{k}}$ with $k_j\geq -D$ for all $1\leq j\leq d$. Thus $f_{\mathbf{i}}\,\Gamma(\mathcal{X},\mathcal{L}^{\otimes D})\subset\Gamma(\widehat{X}_{\mathbf{0}},\mathcal{L}^{\otimes D})$, and we have the injective evaluation map

$$\psi_D:E_D\hookrightarrow F_{\mathbf{Q}},\quad (Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\mapsto\sum_{\mathbf{i}\in V_m^d(\epsilon)}f_{\mathbf{i}}Q_{\mathbf{i}},$$

where $Q_{\mathbf{i}}\in\Gamma(\mathcal{X},\mathcal{L}^{\otimes D})$ and $(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\in E_D$.

Similarly to § 3.1.3, we filter $F_{\mathbf{Q}}$ using the total vanishing order and then the lexicographical ordering within every jet space:

$$F_{\mathbf{Q}}=F_{\mathbf{Q}}^{(\mathbf{0})}\supseteq\cdots\supseteq F_{\mathbf{Q}}^{(\mathbf{n})}\supseteq\cdots,$$

where $\mathbf{n} \in \mathbf{N}^d$, and $F_{\mathbf{Q}}^{(\mathbf{n})} := \mathrm{Span}_{\mathbf{Q}}\{\mathbf{x}^{\mathbf{m}} \,:\, \mathbf{n} \prec \mathbf{m} + D \text{ or } \mathbf{n} = \mathbf{m} + D\}$. Here, the total order $\prec$ on $\mathbf{N}^d$ is defined in § 3.1.3, and for $\mathbf{m} = (m_j)_{j=1}^d$, we define $\mathbf{m} + D := (m_j + D)_{j=1}^d$. The ordering $x_1, \ldots, x_d$ of the variables used to define the lexicographical ordering is immaterial to the proof. In this $\prec$-filtering notation, since $\prod_{j=1}^d x_j^{-D} \in \Gamma(\mathcal{X}, \mathcal{L}^{\otimes D})$ is a generator of $\mathcal{L}_{\mathbf{0}}^{\otimes D}$, where we use $\mathcal{L}_{\mathbf{0}}$ to denote the restriction of $\mathcal{L}$ to the $\mathbf{Z}$-point $\mathbf{x} = \mathbf{0}$, we observe that the $\mathbf{x} = \mathbf{0}$ vanishing order $\mathrm{ord}_{\mathbf{0}}(g(\mathbf{x}))$ of a $g(\mathbf{x}) \in F_{\mathbf{Q}} = \Gamma(\widehat{X}_{\mathbf{0}}, \mathcal{L}^{\otimes D})$ (as a regular section of $\mathcal{L}^{\otimes D}|_{\widehat{X}_{\mathbf{0}}}$, not as a Laurent series in $\mathbf{x}^{-D}\mathbf{Q}[\![\mathbf{x}]\!]$) is at least n if and only if $g(\mathbf{x}) \in F_{\mathbf{Q}}^{(0,\ldots,0,n)}$. We also observe that $g(\mathbf{x}) \in F_{\mathbf{Q}}^{(\mathbf{n})}$ if and only if either $\mathrm{ord}_0(g(\mathbf{x})) > |\mathbf{n}|$ or else $\mathrm{ord}_0(g(\mathbf{x})) = |\mathbf{n}|$ and the lowest lexicographical order term in the homogenous degree $|\mathbf{n}|$ part in g has an exponent vector $\mathbf{m}$ such that $\mathbf{n} \prec \mathbf{m}$. (If here one prefers to think of g as a Laurent series in $\mathbf{x}^{-D}\mathbf{Q}[\![\mathbf{x}]\!]$ rather than as a section of $\mathcal{L}^{\otimes D}|_{\widehat{X}_{\mathbf{0}}}$, one would have to shift all exponent vectors $\mathbf{n}$ to $\mathbf{n} - D$ in these statements.)

As in § 3.1.3, we use $\mathbf{n}^+$ to denote the successor of $\mathbf{n}$ under the total order $\prec$. The graded piece $F_{\mathbf{Q}}^{(\mathbf{n})}/F_{\mathbf{Q}}^{(\mathbf{n}^+)}$ is a one dimensional $\mathbf{Q}$-vector space generated by the image of $\mathbf{x}^{\mathbf{n}-D}$ under the quotient map. The Euclidean lattice structure on $F_{\mathbf{Q}}^{(\mathbf{n})}/F_{\mathbf{Q}}^{(\mathbf{n}^+)}$ is given by the free rank one $\mathbf{Z}$-module generated by the image of $\mathbf{x}^{\mathbf{n}-D}$ and the Euclidean norm with $\|\mathbf{x}^{\mathbf{n}-D}\| = 1$. Note that these Euclidean lattice structures on graded piece are all induced from the free $\mathbf{Z}$-module $F = \mathbf{x}^{-D}\,\mathbf{Z}[\![\mathbf{x}]\!]$ and with the Euclidean norm that has $\{\mathbf{x}^{\mathbf{m}}\}_{\mathbf{m} \in \mathbf{Z}^d_{\geq -D}}$ for an orthonormal basis.

Let $E_D^{(n)} := \psi_D^{-1}\left(F_{\mathbf{Q}}^{(n)}\right) \cap E_D$ denote the preimage of $F_{\mathbf{Q}}^{(\mathbf{n})}$ in E_D under ψ_D. Then ψ_D induces injective maps

$$\psi_D^{(\mathbf{n})} : E_D^{(\mathbf{n})}/E_D^{(\mathbf{n}^+)} \hookrightarrow F_{\mathbf{Q}}^{(\mathbf{n})}/F_{\mathbf{Q}}^{(\mathbf{n}^+)}.$$

into the one-dimensional $\mathbf{Q}$-vector space $F_{\mathbf{Q}}^{(\mathbf{n})}/F_{\mathbf{Q}}^{(\mathbf{n}^+)}$. Therefore $\mathrm{rank}\, E_D^{(\mathbf{n})}/E_D^{(\mathbf{n}^+)} \in \{0, 1\}$ for all $\mathbf{n} \in \mathbf{N}^d$. Let

$$\mathcal{V}_D^d := \{\mathbf{n} \in \mathbf{N}^d \mid \mathrm{rank}\, E_D^{(\mathbf{n})}/E_D^{(\mathbf{n}^+)} = 1\}$$

be the vanishing filtration jumps. We have $\#\mathcal{V}_D^d = \mathrm{rank}\, E_D$.

In the next two subsections § 8.2.11 and § 8.2.29, we provide an upper bound on the local evaluation heights $h_v(\psi_D^{(\mathbf{n})})$ at all $v \in M_{\mathbf{Q}}$. From the definition of the local evaluation height, we need to consider an arbitrary $(Q_{\mathbf{i}})_{\mathbf{i} \in V_m^d(\epsilon)} \in E_D^{(\mathbf{n})} \smallsetminus E_D^{(\mathbf{n}^+)}$ and then provide an upper bound on $\log|c_{\mathbf{n}}|_v - \log\|(Q_{\mathbf{i}})_{\mathbf{i} \in V_m^d(\epsilon)}\|_{E_D, v}$, where $c_{\mathbf{n}}$ denotes the coefficient of $\mathbf{x}^{\mathbf{n}-D}$ in $s := \sum_{\mathbf{i} \in V_m^d(\epsilon)} f_{\mathbf{i}} Q_{\mathbf{i}}$. Here, $|\cdot|_v$ denotes the usual v-adic norm on $\mathbf{Q}$.

Definition 8.2.9. For a formal power series $F(t_1, \ldots, t_d) \in k[\![t_1, \ldots, t_d]\!] = \sum_{\mathbf{n}} a_{\mathbf{n}} \mathbf{t}^{\mathbf{n}}$ over a field k, the n^{th} *order jet* of F is the degree-n homogeneous polynomial $J_n(F) \in k[t_1, \ldots, t_d]_{(n)}$ given by the sum of all the degree-n terms:

$$J_n(F)(t_1, \ldots, t_d) := \sum_{|\mathbf{n}|=n} a_{\mathbf{n}} \mathbf{t}^{\mathbf{n}} \in k[\mathbf{t}]_{(n)}.$$

Remark 8.2.10. In earlier work [Bos01, Bos04] on the algebraization of higher dimensional formal-analytic arithmetic varieties, it sufficed to filter the auxiliary evaluation module using only the total vanishing order at the point $\mathbf{0}$. That we use the finer filtration with one-dimensional quotients has two advantages:

(1) For the estimate of h_v at an archimedean place v, using the product structure of $\overline{\mathbf{D}}^d \to \mathcal{X}(\mathbf{C})$, we have an easy "variable by variable" subharmonic estimate similar to [CDT21, Lemma 2.4.1]. This obviates the blowing-up method in [Bos01, §4.3.2], which gives an upper bound on the Mahler measure of the leading order jet polynomial $J_n(F) \in \mathbf{Q}[\mathbf{x}]_{(n)}$, whereas the quantities that need to be estimated are the individual coefficients $c_{\mathbf{n}}$ in that leading order jet. The discrepancy in these quantities is too sensitive in the dimension $d = \dim \mathcal{X}_{\mathbf{Q}}$, which is fixed in [Bos01] whereas we want to have $d \to \infty$ at the end.

(2) The advantage for the h_v estimate at a non-archimedean place v is crucial. Among all $\mathbf{n}$ with $|\mathbf{n}| = n$, due to our specific construction of E_D, we will have a much better estimate of $h_v\left(\psi_D^{(\mathbf{n})}\right)$ under the condition that $\{n_j\}_{j=1}^d$ has asymptotically equidistributed components. Our complete filtration with including the lexicographical ordering allows us to take stock of this improvement. △

8.2.11. *Archimedean estimate.* Recall that by the same reduction argument as in the beginning of the proof of Lemma 7.4.1, we may assume that $f_i(\varphi(z))$ is meromorphic on an open neighborhood of $|z| \leq 1$, for all $1 \leq i \leq m$.

For ease of notation, we use $|\cdot|$ to denote the usual absolute value $|\cdot|_\infty$. The broader outline of the proof is similar to [Bos01, §4.3.3], with caveat the modification we described in Remark 8.2.10(1). Given $s \in \psi_D^{(\mathbf{n})}(E_D^{\mathbf{n}} \smallsetminus E_D^{\mathbf{n}^+})$, we will follow [CDT21, § 2.4] in studying the n^{th} (leading) order jet $J_n(\varphi^* s)(\mathbf{z}) \in \mathbf{C}[\mathbf{z}]_{(n)}$ at the point $\mathbf{z} = \mathbf{0} \in \overline{\mathbf{D}}^d$. Here and in the following, we use the notation $n := |\mathbf{n}|$, and by a slight abuse of notation, we continue to denote $\varphi : \overline{\mathbf{D}}^d \to \mathcal{X}(\mathbf{C})$ for the analytic morphism $\mathbf{x} \mapsto (\varphi(z_1), \dots, \varphi(z_d))$ given by $\varphi : \overline{\mathbf{D}} \to \mathbf{C}$ diagonally on each factor. By extension of that notation, and in a manner unifying § 6 and § 7.4, we consider for every $\mathbf{r} \in (0,1]^d$ the analytic morphism

$$\varphi_{\mathbf{r}} : \overline{\mathbf{D}}^d \to \mathbf{C}^d \hookrightarrow \mathcal{X}(\mathbf{C}), \qquad \mathbf{z} \mapsto (\varphi(r_1 z_1), \dots, \varphi(r_d z_d)).$$

We will use the Poisson–Jensen formula to bound $\log |c_{\mathbf{n}}|$ in terms of the jet function, then relate the resulting bound to the Chern form of $\overline{\mathcal{L}}$ by means of the Poincaré–Lelong formula. We follow the notations in § 7.4, and we borrow from (7.1.7) the notation for the slopes α_k. In addition, for each $\mathbf{n} \in \mathbf{N}^d$, we define $r(\mathbf{n}) := (r(n_1), \dots, r(n_d))$, where

$$r(t) := \begin{cases} r_k, & \text{if } t/D \in [\alpha_k, \alpha_{k+1}), \\ 1, & \text{if } t/D > m. \end{cases} \tag{8.2.12}$$

Throughout this section, $\mathbf{z} = (z_1, \dots, z_d)$ denotes the coordinate on $\overline{\mathbf{D}}^d$. We use the trivialization $\mathcal{L}^{\otimes D}|_{\mathbf{C}^d} \xrightarrow{\simeq} \mathcal{O}_{\mathbf{C}^d}$ of $\mathcal{L}^{\otimes D}$ over $\mathbf{A}^d_{\mathbf{C}}$ given by $\mathbf{x}^{-D} \mapsto \mathbf{1}$. Under this identification, $s/\mathbf{x}^{-D} = s \cdot \mathbf{x}^D =: G(\mathbf{x})$ is naturally in $\Gamma(\widehat{X}_{\mathbf{0}}, \mathcal{O}_{\widehat{X}_{\mathbf{0}}})$ with vanishing order n. Since our analytic $\varphi_{r(\mathbf{n})}$-pullback is defined by $x_j = \varphi_{r(n_j)}(z_j) = (\varphi'(0) \cdot r(n_j)) z_j + O(z_j^2)$, we have by construction that $\varphi^*_{r(\mathbf{n})} G$ has $\mathbf{z} = \mathbf{0}$ vanishing order n, with $c_{\mathbf{n}} \varphi'(0)^n \prod_{j=1}^d r(n_j)\, \mathbf{z}^{\mathbf{n}}$ for the lexicographically minimal term in $J_n(\varphi^*_{r(\mathbf{n})}G)$ (as well as the overall $\prec$-minimal term in $\varphi^*_{r(\mathbf{n})}G$). However, since we only assume in this theorem the *meromorphy* (as opposed to the holomorphy) of the pullbacks: $\varphi^* f_i \in \mathcal{M}(\overline{\mathbf{D}})$, the analytic germ $\varphi^*_{r(\mathbf{n})}G \in \mathbf{C}[\![\mathbf{z}]\!]$ only extends meromorphically, rather than holomorphically through $\overline{\mathbf{D}}$. But if we choose $h \in \mathcal{O}(\overline{\mathbf{D}})$ a holomorphic function such that $h(0) = 1$ and all $h \cdot \varphi^* f_i \in \mathcal{O}(\overline{\mathbf{D}})$ are holomorphic, then $h(z_1)\cdots h(z_d)\cdot \left(\varphi^*_{r(\mathbf{n})}\right) G(\mathbf{z}) \in \mathcal{O}\left(\overline{\mathbf{D}}^d\right)$ is holomorphic throughout $\overline{\mathbf{D}}^d$, and has the same leading order jet $J_n\left(\varphi^*_{r(\mathbf{n})}G\right)$ as $\varphi^*_{r(\mathbf{n})}G$.

This puts us in a position to use [CDT21, Lemma 2.4.1] for upper-bounding the requisite coefficient $|c_{\mathbf{n}}|$ in terms of the Mahler measure of the (homogeneous) polynomial $J_n\left(\varphi^*_{r(\mathbf{n})}G\right) = J_n\left(h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right)$, in which certainly the overall lexicographically lowest term is the monomial $c_{\mathbf{n}}\mathbf{z}^{\mathbf{n}}$ in the multidegree $\mathbf{n}$:

$$\log|c_{\mathbf{n}}| \leq -n\log|\varphi'(0)| - \sum_{j=1}^d n_j \log r(n_j) + \int_{\mathbf{T}^d} \log\left|J_n\left(h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right)\right| \mu_{\mathrm{Haar}}. \tag{8.2.13}$$

We can connect this to Bost's blowing up argument in [Bos01, Equation (4.28) in Lemma 4.13] which shows

$$\begin{aligned}\int_{\mathbf{T}^d} \log\left|J_n\left(h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right)\right| \mu_{\mathrm{Haar}} &\leq \int_{\mathbf{T}^d}\log\left|h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right| \mu_{\mathrm{Haar}} \\ &= \int_{\mathbf{T}^d}\log\left|\varphi^*_{r(\mathbf{n})}G\right| \mu_{\mathrm{Haar}} + O_h(1).\end{aligned} \tag{8.2.14}$$

To recall Bost's argument, write $h(z_1)\cdots h(z_d)\cdot G(\varphi^*_{r(\mathbf{n})}(\mathbf{z})) =: \sum_{\mathbf{k}\in\mathbf{N}^d} c'_{\mathbf{k}}\mathbf{z}^{\mathbf{k}}$, and define

$$U_t(\mathbf{z}) := \sum_{\mathbf{k}\in\mathbf{N}^d} c'_{\mathbf{k}} t^{|\mathbf{k}|-n}\mathbf{z}^{\mathbf{k}}, \qquad \text{for } t\in\mathbf{C} \text{ with } |t|\leq 1. \tag{8.2.15}$$

Then, by the $\mathbf{z}\mapsto t\mathbf{z}$ substitution,

$$\int_{\mathbf{T}^d}\log|U_t(\mathbf{z})|\,\mu_{\mathrm{Haar}}(\mathbf{z}) = \int_{|t|\mathbf{T}^d}\log\left|h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right| \mu_{\mathrm{Haar}}(\mathbf{z}) - n\log|t|. \tag{8.2.16}$$

In particular, $\int_{\mathbf{T}^d}\log|U_t(\mathbf{z})|\,\mu_{\mathrm{Haar}}(\mathbf{z})$, which by (8.2.15) is clearly a subharmonic function in the single complex variable t, only depends on that complex variable t through $|t|$ by means of the right-hand side of the identity (8.2.16). Hence the value of that function at $t=1$, which equals $\int_{\mathbf{T}^d}\log\left|h(z_1)\cdots h(z_d)\cdot\varphi^*_{r(\mathbf{n})}G\right| \mu_{\mathrm{Haar}}$ but is also the $t\in\mathbf{T}$ integral of $\int_{\mathbf{T}^d}\log|U_t(\mathbf{z})|\,\mu_{\mathrm{Haar}}(\mathbf{z})$, is no less than its value at $t=0$, which is the left-hand side of the requisite bound (8.2.14).

Next we follow the proof idea of [Bos20, Theorem 10.5.3], along with the discussion in [BC22, §§4.2–4.3] in order to pass from smooth Green functions to $\mathcal{C}^{\mathrm{b}\Delta}$ ones. By the product structure of $\overline{\mathcal{L}}$ and the computation from the proof of Lemma 7.4.1, we have

$$\int_{\mathbf{T}^d} \log \|\varphi_{r(\mathbf{n})}^* \mathbf{x}^{-D}\|_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\, \mu_{\mathrm{Haar}} = \sum_{j=1}^{d} \int_{\mathbf{T}} \log \|\varphi_{r(n_j)}^*(x_j^{-D})\|_{\overline{\mathcal{O}(D)}}\, \mu_{\mathrm{Haar}}$$
$$= \sum_{j=1}^{d} \left(- \int_{\overline{\mathbf{D}}} \log^+ |z|^{-1}\, c_1\left(\varphi_{r(n_j)}^*\overline{\mathcal{O}(D)}\right) + \|\varphi_{r(n_j)}^*(x^{-D})\|_{\varphi_{r(n_j)}^*\overline{\mathcal{O}(D)}} \,|_{z=0} \right). \tag{8.2.17}$$

Since $x^{-D}\,|_{x=0}$ is a $\mathbf{Z}$-generator of the free $\mathbf{Z}$-module $\mathcal{O}(D)_0$, we have

$$\|\varphi_{r(n_j)}^*(x^{-D})\|_{\overline{\mathcal{O}(D)}}\,|_{z=0} = -\widehat{\deg}\,\overline{\mathcal{O}(D)}_0 = -D\,\widehat{\deg}\,\overline{\mathcal{O}(1)}_0.$$

Putting together (8.2.13), (8.2.14), and (8.2.17), we have

$$\begin{aligned}
&\log |c_{\mathbf{n}}| - \int_{\mathbf{T}^d} \log \|\varphi_{r(\mathbf{n})}^* s\|_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\, \mu_{\mathrm{Haar}} \\
&\le -n \log |\varphi'(0)| - \sum_{j=1}^{d} n_j \log r(n_j) \\
&\quad + \int_{\mathbf{T}^d} \left(\log |\varphi_{r(\mathbf{n})}^* G(\mathbf{z})| - \log \|\varphi_{r(\mathbf{n})}^* s\|_{\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}}} \right) \mu_{\mathrm{Haar}} + O_h(1) \\
&= -n \log |\varphi'(0)| - \sum_{j=1}^{d} n_j \log r(n_j) - \int_{\mathbf{T}^d} \log \|\varphi_{r(\mathbf{n})}^* \mathbf{x}^{-D}\|_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\, \mu_{\mathrm{Haar}} + O_h(1) \\
&= -n \log |\varphi'(0)| - \sum_{j=1}^{d} n_j \log r(n_j) \\
&\qquad + D \left(\sum_{j=1}^{d} \widehat{\deg}\,(\overline{\mathcal{O}(1)}_0) + \int_{\overline{\mathbf{D}}} \log^+ |z|^{-1} c_1(\varphi_{r(n_j)}^*\overline{\mathcal{O}(1)}) \right) + O_h(1) \\
&= -n \log |\varphi'(0)| - \sum_{j=1}^{d} n_j \log r(n_j) + D \sum_{j=1}^{d} \widehat{T}(r(n_j),\varphi) + O_h(1),
\end{aligned} \tag{8.2.18}$$

where $\widehat{T}(r(n_j),\varphi))$ is the Bost–Charles characteristic we defined in 7.1.2, and the final equality derives from the projection formula [BC22, Proposition 7.2.2] applied to the morphism $(\iota, \varphi_{r(n_j)})$ in Remark 7.3.2 and Lemma 7.4.5. We spell out that last step in more detail. Following the definitions of pullback of Hermitian vector bundles in[BC22, §7.1.1.1] and the arithmetic intersection number in [BC22, Equation 6.2.4], we have

$$(\iota, \varphi_{r(n_j)})^*\overline{\mathcal{O}(1)} \cdot ([0], \log^+ |z|^{-1}) = \widehat{\deg}\,\overline{\mathcal{O}(1)}_0 + \int_{\overline{\mathbf{D}}} \log^+ |z|^{-1} c_1(\varphi^*(\overline{\mathcal{O}(1)})).$$

By the projection formula [BC22, Proposition 7.2.2] together with Lemma 7.4.5, we derive the requisite equality

$$(\iota, \varphi_{r(n_j)})^*\overline{\mathcal{O}(1)} \cdot ([0], \log^+ |z|^{-1}) = \widehat{T}(r(n_j),\varphi)),$$

and the final line on (8.2.18) follows.

We claim that (8.2.18) implies for an arbitrary $(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)} \in E_D^{(\mathbf{n})} \smallsetminus E_D^{(\mathbf{n}^+)}$ a uniform upper bound:

$$\log|c_{\mathbf{n}}|-\|(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\| \leq -n\log|\varphi'(0)|-\sum_{j=1}^d n_j\log r(n_j)+D\sum_{j=1}^d \widehat{T}(r(n_j),\varphi))+o(D). \tag{8.2.19}$$

As μ_{Haar} is not a continuous measure on $\overline{\mathbf{D}}^d$, we begin by approximating $\log^+|z|^{-1}$ by a sequence $(g_k)_{k\in\mathbf{N}}$ of smooth rotationally symmetric Green functions on $\overline{\mathbf{D}}$ for the divisor $[0]$. Precisely, by [Bos20, page 268], we choose $g_k \in C^\infty(\overline{\mathbf{D}})$ with $\operatorname{supp}(g_k)\subset \mathbf{D}$, such that $g_k(z)=g_k(|z|)$ and $g_k-\log^+|z|^{-1}\to 0$ uniformly on $\overline{\mathbf{D}}$. Following the same argument as in the proof of Lemma 7.4.1 (see, for instance, [Bos20, pages. 270–271] for the one-dimensional case), we write $\frac{i}{\pi}\partial\overline{\partial}g_k=-\delta_0+\mu_k$, where μ_k is a smooth probability measure on $\overline{\mathbf{D}}$, and then, denoting by μ_k^d the product measure induced from μ_k on $\overline{\mathbf{D}}^d$:

$$\begin{aligned}\int_{\overline{\mathbf{D}}^d}\log\|\varphi_{r(\mathbf{n})}^*\mathbf{x}^{-D}\|_{\varphi_{(r(\mathbf{n})}^*\overline{\mathcal{L}})}\,\mu_k^d &= -D\sum_{j=1}^d\left(\int_{\overline{\mathbf{D}}} g_k c_1(\varphi_{r(n_j)}^*\overline{\mathcal{O}(1)})+\widehat{\deg}\,\overline{\mathcal{O}(1)}_0\right)\\ &= -D\sum_{j=1}^d(\iota,\varphi_{r(n_j)})^*\overline{\mathcal{O}(1)}\cdot([0],g_k).\end{aligned}$$

Moreover, since g_k and μ_k are rotationally invariant, Poisson–Jensen gives:

$$\begin{aligned}\log|c_{\mathbf{n}}|+n\log|\varphi'(0)|+\sum_{j=1}^d n_j\log r(n_j) &\leq \int_{\overline{\mathbf{D}}^d}\left(\log|J_n(\varphi_{r(\mathbf{n})}^*G)(\mathbf{z})|-\log|\mathbf{z}|^{\mathbf{n}}\right)\mu_k^d\\ &= \int_{\overline{\mathbf{D}}^d}\log|J_n(\varphi_{r(\mathbf{n})}^*G)|\,\mu_k^d-n\int_{\overline{\mathbf{D}}}\log|z|\mu_k;\end{aligned}$$

and our previous argument gives

$$\int_{\overline{\mathbf{D}}^d}\log|J_n(\varphi_{r(\mathbf{n})}^*G)|\,\mu_k^d\leq\int_{\overline{\mathbf{D}}^d}\log|\varphi_{r(\mathbf{n})}^*G|\,\mu_k^d.$$

Since μ_k is smooth on $\overline{\mathbf{D}}^d$ and $\overline{\mathbf{D}}^d$ is compact, there exists a constant $C_k>0$ depending only on $d, g_k, \varphi, \mathbf{r}$ (but independent of n, D, and s), such that $C_k\varphi_{r(\mathbf{n})}^*\nu = C_k\bigwedge_{j=1}^d\mathrm{pr}_j^*\varphi_{r(n_j)}^*\,\omega_{\mathrm{FS}}\geq\mu_k^d$ as a pointwise inequality for smooth (d,d)-forms on $\overline{\mathbf{D}}^d$.

Therefore, we have

$$
\begin{aligned}
&\log \|(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\| \\
&\geq \max_{\mathbf{i}\in V_m^d(\epsilon)} \log\|Q_{\mathbf{i}}\| = \frac{1}{2}\max_{\mathbf{i}\in V_m^d(\epsilon)} \log \int_{\mathcal{X}(\mathbf{C})} \|Q_{\mathbf{i}}\|^2_{\overline{\mathcal{L}}}\,\nu \\
&\geq -\frac{1}{2}\log C_k + \frac{1}{2}\max_{\mathbf{i}\in V_m^d(\epsilon)} \log \int_{\overline{\mathbf{D}}^d} \|\varphi_{r(\mathbf{n})}^* Q_{\mathbf{i}}\|^2_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\,\mu_k^d \\
&\geq -\frac{1}{2}\log C_k - \log\left\{ m^d(D+1)^d \max_{\substack{1\leq i\leq m,\\ z\in\overline{\mathbf{D}}}} |f_i(z)|^d \right\} + \frac{1}{2}\log\int_{\mathbf{D}^d} \|\varphi_{r(\mathbf{n})}^* s\|^2_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\,\mu_k^d \\
&> -C_k' - d\log D + \frac{1}{2}\log\int_{\mathbf{D}^d} \|\varphi_{r(\mathbf{n})}^* s\|^2_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\,\mu_k^d \\
&\geq -C_k' - d\log D + \int_{\mathbf{D}^d} \log\|\varphi_{r(\mathbf{n})}^* s\|_{(\varphi_{r(\mathbf{n})}^*\overline{\mathcal{L}})}\,\mu_k^d,
\end{aligned}
$$

where $C_k' > 0$ is a constant only depending on $d, m, \{f_i\}, g_k, \varphi, \mathbf{r}$ (but independent of n, D, and s), and the last inequality follows from the quadratic mean — geometric mean inequality since μ_k^d is a probability measure on $\overline{\mathbf{D}}^d$ (see for instance [BGS94, (1.4.10)]).

We get:

$$
\begin{aligned}
\log|c_{\mathbf{n}}| - \log\|(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\| \leq\ & -n\log|\varphi'(0)| - \sum_{j=1}^d n_j \log r(n_j) - n\int_{\overline{\mathbf{D}}} \log|z|\,\mu_k \\
& + D\sum_{j=1}^d (\iota, \varphi_{r(n_j)})^* \overline{\mathcal{O}(1)}\cdot([0], g_k) + o(D),
\end{aligned}
\tag{8.2.20}
$$

giving for any fixed $\gamma \geq 0$ independent of k and D, and for all large enough $k \gg 1$:

$$
\begin{aligned}
\limsup_{\substack{D\to\infty\\ |\mathbf{n}|\geq\gamma D}}{}' \frac{h_\infty(\psi_D^{(\mathbf{n})}) + \sum_{j=1}^d n_j\log r(n_j)}{D} \leq & \sum_{j=1}^d (\iota, \varphi_{r(n_j)})^* \overline{\mathcal{O}(1)}\cdot([0], g_k) \\
& - \gamma\left(\log|\varphi'(0)| + \int_{\overline{\mathbf{D}}} \log|z|\,\mu_k\right),
\end{aligned}
\tag{8.2.21}
$$

where the dash over the limit supremum indicates that we consider all $\mathbf{n} \in \mathcal{V}_D^d$ with $|\mathbf{n}| \geq \gamma D$ and sharing some fixed $r(\mathbf{n})$ (note that there are only $(l+1)^d$ many possibilities of $r(\mathbf{n})$). Here, remarking that $\log|\varphi'(0)| > 0$ while the uniform limit $g_k \to \log^+|z|^{-1}$ on $\overline{\mathbf{D}}$ implies

$$\int_{\overline{\mathbf{D}}} \log|z|\,\mu_k \to 0,$$

the meaning of "large enough k" is specified by the positivity of the term $\log|\varphi'(0)|+\int_{\overline{\mathbf{D}}}\log|z|\,\mu_k$. At this point, the requisite estimate (8.2.19) follows[34] form the convergence

$$\lim_{k\to\infty}\Big((\iota,\varphi_{r(\mathbf{n})})^*\overline{\mathcal{O}(1)}\cdot([0],g_k)\Big)=\Big(\overline{\mathcal{O}(1)}\cdot(\iota,\varphi_{r(\mathbf{n})})_*([0],\log^+|z|^{-1})\Big)$$
$$=\sum_{j=1}^{d}\widehat{T}(r(n_j),\varphi).$$

Armed with (8.2.19), and using inputs from the functional bad approximability theorems in § 3.2 (which are in place since, in all cases, the f_i are holonomic functions; noting that the proof of Lemma 7.3.17 entails automatic holonomicity from the condition $\log|\varphi'(0)|>\sum_{h=1}^r\max_{1\le i\le m}b_{i,h}$), we now estimate the total contribution

$$\sum_{\mathbf{n}\in\mathbf{N}^d}\operatorname{rank}\Big(E^{(\mathbf{n})}/E^{(\mathbf{n}^+)}\Big)\cdot h_\infty(\psi_D^{(\mathbf{n})})=\sum_{\mathbf{n}\in\mathcal{V}_D^d}h_\infty(\psi_D^{(\mathbf{n})})$$

of the archimedean height showing in the right-hand side of Bost's slopes inequality (7.2.14).

For any $\epsilon'>0$, Lemma 3.2.14 shows that all $D\gg_{\epsilon',\{f_i\}}1$ satisfy

$$\mathcal{V}_D^d\subset[0,(m+\epsilon')D]^d.$$

(In the Lemma, we may pick $\varepsilon:=\epsilon'/2$ and we consider $D\gg_{\epsilon',\{f_i\}}1$ such that $\epsilon' D/2>C(\varepsilon)$.) By (8.2.19), which by definition is an upper bound on the $\mathbf{n}^{\text{th}}$ archimedean evaluation height $h_\infty(\psi_D^{\mathbf{n}})$, we have:

$$\sum_{\mathbf{n}\in\mathcal{V}_D^d}h_\infty(\psi_D^{(\mathbf{n})})$$
$$\le-\log|\varphi'(0)|\left(\sum_{\mathbf{n}\in\mathcal{V}_D^d}|\mathbf{n}|\right)+D\sum_{\mathbf{n}\in\mathcal{V}_D^d}\sum_{j=1}^{d}\Big(-\frac{n_j}{D}\log r(n_j)+\widehat{T}(r(n_j),\varphi)\Big)\Big)+o(D^{d+1}).\tag{8.2.22}$$

[34]To be fully rigorous, the proof of (8.2.19) is completed by the paragraph below. We firstly note a mild sloppiness in our formulation due to the involvement of γ and the requirement to only work with the $\mathbf{n}$ constrained by $|\mathbf{n}|/D\ge\gamma$. In practical terms, we use a large deviations bound to show that $|\mathbf{n}|/D=d(m/2+o(1))$ for most $\mathbf{n}\in\mathcal{V}_D^d$, and take $\gamma=d(m/2-\epsilon_0)$ with $\epsilon_0\to0$ in the end. Then (8.2.21) is used for these $\mathbf{n}$, while a trivial estimate applies to the leftover meagre set of $\mathbf{n}$.

In any case, here is a rigorous completion of the proof of the requisite bound (8.2.19). For any $\epsilon>0$, we can pick $k\gg1$ (depending on d) such that $\int_{\overline{\mathbf{D}}}\log|z|\,\mu_k<\epsilon$, and $\Big|(\iota,\varphi_{r(\mathbf{n})})^*\Big(\overline{\mathcal{O}(1)}\cdot([0],g_k)\Big)-\sum_{j=1}^d\widehat{T}(r(n_j),\varphi))\Big|<\epsilon$. For this specific k, we consider $D\gg1$ such that the $o(D)$ in (8.2.20) is $<\epsilon D/2$. From (8.2.21),

$$h_\infty(\psi_D^{(\mathbf{n})})\le-n\log|\varphi'(0)|-\sum_{j=1}^{d}n_j\log r(n_j)+D\sum_{j=1}^{d}\widehat{T}(r(n_j),\varphi))+\epsilon(n+D).$$

At this point (8.2.19) follows by Lemma 7.3.17 and Shidlovsky's lemma, which give $n=|\mathbf{n}|=O(dD)$ for $\mathbf{n}\in\mathcal{V}_D^d$.

We estimate the quantities

$$\sum_{\mathbf{n}\in\mathcal{V}_D^d}|\mathbf{n}|,\ \sum_{\mathbf{n}\in\mathcal{V}_D^d}\sum_{j=1}^{d}(n_j/D)\log r(n_j),\ \sum_{\mathbf{n}\in\mathcal{V}_D^d}\sum_{j=1}^{d}\widehat{T}(r(n_j),\varphi))$$

using the following consequence of Theorem 4.2.1, to the effect that most $\mathbf{n}\in\mathcal{V}_D^d$ have uniformly distributed components. Similarly to § 6.2, recall from the statement of Theorem 8.0.1, we use $P_\varepsilon^d(N)\subset[0,N]^d\cap\mathbf{Z}^d$ to denote the subset of those $\mathbf{n}$ for which the normalized $([0,1),\mu_{\text{Lebesgue}})$ discrepancy of $\{n_i/N\}_{i=1}^d$ is $\le\varepsilon$, and $B_d^\varepsilon(N)$ to denote the complement of $P_\varepsilon^d(N)$ in $[0,N]^d\cap\mathbf{Z}^d$.

Lemma 8.2.23. *There is a function* $c:(0,1)\to(0,1)$*, such that the following holds.*

Consider an $\epsilon''\in(0,1)$*. Then, for all* $\epsilon'>0$ *small enough with respect to* ϵ'',

$$\lim_{D\to\infty}\frac{\#\left\{\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)\right\}}{\#\mathcal{V}_D^d}=O\left(e^{-c(\epsilon'')d}\right),$$

where the implicit coefficient is absolute.

Proof. Note that $\#\left\{\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)\right\}\le\#B_d^{\epsilon''}((m+\epsilon')D)$, and recall that

$$\lim_{d\to\infty}\lim_{D\to\infty}\frac{\#\mathcal{V}_D^d}{m^dD^d}=\frac{\operatorname{rank}E_D}{m^dD^d}=1.$$

Hence, it suffices to show that $\lim_{D\to\infty}\frac{\#B_d^{\epsilon''}((m+\epsilon')D)}{m^dD^d}=O(e^{-cd})$ for some $c=c(\epsilon'')>0$ and a small enough ϵ'.

By Theorem 4.2.1, we have

$$\lim_{D\to\infty}\#B_d^{\epsilon''}((m+\epsilon')D)/((m+\epsilon')D)^d=O(e^{-c_0(\epsilon'')d}),$$

with a certain $c_0(\epsilon'')>0$ and an absolute implicit coefficient. For ϵ' sufficiently small in terms of $c_0(\epsilon'')$, we will have $\lim_{D\to\infty}((m+\epsilon')D)^d/(mD)^d=O(e^{c_0(\varepsilon'')d/2})$. We obtain the desired bound with $c:=c_0/2$. □

Now, for an arbitrary $\epsilon''>0$, we pick the sufficiently small $\epsilon'>0$ as guaranteed by Lemma 8.2.23, and apply Lemma 3.2.14 as discussed above to obtain that $\mathcal{V}_D^d\subset[0,(m+\epsilon')D]^d$ for $D\gg1$. We obtain:

$$\begin{aligned}\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d}|\mathbf{n}|}{dm^dD^{d+1}}\right\}&\ge\lim_{d\to\infty}\lim_{D\to\infty}\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)}|\mathbf{n}|}{dD\#\mathcal{V}_D^d}\\&\ge\lim_{d\to\infty}\lim_{D\to\infty}\frac{(\#\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D))(m+\epsilon')(1-2\sqrt{\epsilon''})/2}{\#\mathcal{V}_D^d}\\&=(m+\epsilon')(1-2\sqrt{\epsilon''})/2.\end{aligned}$$

Here, the second inequality follows upon remarking that the definition of the discrepancy function implies $|\mathbf{n}|\ge dD(m+\epsilon')(1-2\sqrt{\epsilon''})/2$ for all $\mathbf{n}\in P_d^{\epsilon''}((m+\epsilon')D)$; and the last equality follows from Lemma 8.2.23 which implies that, for fixed ϵ'',ϵ',

$$\lim_{d\to\infty}\lim_{D\to\infty}\frac{\#\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)}{\#\mathcal{V}_D^d}=\lim_{d\to\infty}\left\{1-O(e^{-c(\epsilon'')d})\right\}=1.$$

Now we let $\epsilon'' \to 0$ (this will force $\epsilon' \to 0$). We get:

$$\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d}|\mathbf{n}|}{dm^dD^{d+1}}\right\} \geq \frac{m}{2}. \tag{8.2.24}$$

Similarly, still directly from the definition of the discrepancy function, we have the following evaluation for all $\mathbf{n}\in\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)$:

$$\begin{aligned}&\sum_{j=1}^d\left\{-(n_j/D)\log r(n_j)+\widehat{T}\left(r(n_j),\varphi\right)\right\}\\&=\frac{d}{m}\sum_{k=0}^l\left\{(\alpha_{k+1}-\alpha_k)\widehat{T}\left(r_k,\varphi\right)-\frac{1}{2}(\alpha_{k+1}^2-\alpha_k^2)\log r_k\right\}+O\left(d(\epsilon'+\sqrt{\epsilon''})\right),\end{aligned} \tag{8.2.25}$$

recalling the notations (7.1.7) and (8.2.12). (The implicit coefficient here depends linearly on $m|\log r_0|+\widehat{T}(1,\varphi)$.) A partial summation now gives

$$\begin{aligned}&\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)}\sum_{j=1}^d\left(-(n_j/D)\log r(n_j)+\widehat{T}(r(n_j),\varphi)\right)}{dm^dD^d}\right\}\\&\quad=\widehat{T}(1,\varphi)-\frac{1}{2m}\sum_{k=1}^l\frac{(\widehat{T}(r_k,\varphi)-\widehat{T}(r_{k-1},\varphi))^2}{\log r_k-\log r_{k-1}}+O((\epsilon'+\sqrt{\epsilon''})),\end{aligned} \tag{8.2.26}$$

and the last error term goes to 0 once we take $\epsilon''\to 0$ (which implies $\epsilon'\to 0$).

Further, for all $\mathbf{n}\in\mathcal{V}_D^d$, we certainly have the following trivial bound

$$\sum_{j=1}^d\left(-(n_j/D)\log r(n_j)+\widehat{T}\left(r(n_j),\varphi\right)\right)\leq d(m+\epsilon')|\log r_0|+d\widehat{T}(1,\varphi),$$

and thus

$$\lim_{d\to\infty}\lim_{D\to\infty}\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)}\sum_{j=1}^d\left(-(n_j/D)\log r(n_j)+\widehat{T}(r(n_j),\varphi)\right)}{dm^dD^d}=0. \tag{8.2.27}$$

Combining (8.2.22), (8.2.24), (8.2.26), and (8.2.27), we arrive at our total archimedean evaluation height bound:

$$\begin{aligned}&\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d}h_\infty(\psi_D^{(\mathbf{n})})}{dm^dD^{d+1}}\right\}\\&\quad\leq-\frac{m}{2}\log|\varphi'(0)|+\widehat{T}(1,\varphi)-\frac{1}{2m}\sum_{k=1}^l\frac{(\widehat{T}(r_k,\varphi)-\widehat{T}(r_{k-1},\varphi))^2}{\log r_k-\log r_{k-1}}.\end{aligned} \tag{8.2.28}$$

8.2.29. *Non-archimedean estimate.* The main idea here is similar to § 6.6.

Let $h_{\mathrm{fin}}(\psi_D^{(\mathbf{n})})$ denote $\sum_{v\nmid\infty}h_v(\psi_D^{(\mathbf{n})})$. For $\mathbf{n}\in\mathcal{V}_D^d$ and a prime p, by definition, $h_p(\psi_D^{(\mathbf{n})})$ is $\log p$ times the maximal p-adic valuation $v_{p,\mathbf{n}}$ of the denominators of the $(\mathbf{n}-D)$-th coefficient of $\sum_{\mathbf{i}\in V_m^d(\epsilon)}f_{\mathbf{i}}Q_{\mathbf{i}}$ across all $(Q_{\mathbf{i}})_{\mathbf{i}\in V_m^d(\epsilon)}\in E_D$. Since all $Q_{\mathbf{i}}(\mathbf{x})$ are $\mathbf{Z}$-linear combinations of monomials $\mathbf{x}^{\mathbf{k}}$ with $\mathbf{k}\in[-D,0]^d$, it follows that $v_{p,\mathbf{n}}$ is at most the maximum of the p-adic valuations of the denominators of the

$\mathbf{x}^{\mathbf{m}}$ coefficients of all $f_{\mathbf{i}}(\mathbf{x})$ for all $\mathbf{m}$ with $(n_j - D)^+ \le m_j \le n_j$ for all $1 \le j \le d$; here for once we write $(n_j - D)^+ := \max(n_j - D, 0)$. We consider separately the $\prod_{h=1}^r [1,\dots,b_{i,h}n]$ and the n^{e_i} pieces of the p-denominators of the coefficients of $f_i(x)$:

$$v_{p,\mathbf{n}}^{\flat} := \max_{\mathbf{m}\preceq\mathbf{n},\, \mathbf{i}\in V_m^d(\epsilon)} \left\{ \sum_{j=1}^d \mathrm{val}_p([1,\dots,b_{i_j,1}\cdot m_j]\cdots[1,\dots,b_{i_j,r}\cdot m_j]) \right\},$$

$$v_{p,\mathbf{n}}^{\sharp} := \max_{\substack{\mathbf{i}\in V_m^d(\epsilon)\\ (n_j-D)^+\le m_j\le n_j, \forall 1\le j\le d}} \left\{ \sum_{j=1}^d e_{i_j}\,\mathrm{val}_p(\max\{m_j,1\}) \right\},$$

where val_p denotes the usual p-adic valuation with $\mathrm{val}_p(p) = 1$. Here we use the convention that for $m_j = 0$, we set $[1,\dots,b_{i_j,h}\cdot m_j] = 1$. By definition, we have

$$v_{p,\mathbf{n}} \le v_{p,\mathbf{n}}^{\flat} + v_{p,\mathbf{n}}^{\sharp}. \tag{8.2.30}$$

We continue with the notations from § 8.2.11; in particular, $\mathcal{V}_D^d \subset [0,(m+\epsilon')D]^d$ by Lemma 3.2.14. We firstly discuss $v_{p,\mathbf{n}}^{\flat}$. For the case $\mathbf{n} \in \mathcal{V}_D^d \cap B_d^{\epsilon''}((m+\epsilon')D)$, we stick to the trivial bound. Observe that $\prod_{h=1}^r [1,\dots,(\max_{1\le i\le m} b_{i,h})\cdot n]$ is a multiple of the denominators of the x^n-coefficients of all $f_1,\dots,f_m$. By the prime number theorem, it follows that

$$\sum_p v_{p,\mathbf{n}}^{\flat}\log p \le \left(\sum_{h=1}^r \max_{1\le i\le m} b_{i,h}\right)|\mathbf{n}| + o(|\mathbf{n}|). \tag{8.2.31}$$

Summing over all $\mathbf{n} \in \mathcal{V}_D^d \cap B_d^{\epsilon''}((m+\epsilon')D)$, so that in particular $|\mathbf{n}| \le d(m+\epsilon')D$, we have, as $d\to\infty$:

$$\begin{aligned} &\limsup_{D\to\infty}\left\{ \frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)}\sum_p v_{p,\mathbf{n}}^{\flat}\log p}{dm^dD^{d+1}} \right\} \\ &\le \limsup_{D\to\infty}\left\{ \frac{\left(\sum_{h=1}^r \max_{1\le i\le m} b_{i,h}\right)(m+\epsilon')Dd}{dD}\,\frac{\#\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)}{m^dD^d} \right\} \\ &= O\left(\left(\sum_{h=1}^r \max_{1\le i\le m} b_{i,h}\right)(m+\epsilon')e^{-c(\epsilon'')d} \right) = o_{d\to\infty}(1). \end{aligned} \tag{8.2.32}$$

For the case $\mathbf{n} \in \mathcal{V}_D^d \cap P_d^{\epsilon''}((m+\epsilon')D)$: at a fixed p, our assumption on the denominator types of the f_i says that the p-adic valuation of the denominators of the coefficients of all the monomials with exponent vectors $\preceq \mathbf{n}$ in $f_{\mathbf{i}}$ is at most

$$\sum_{j=1}^d \mathrm{val}_p\left([1,\dots,b_{i_j,1}\cdot n_j]\cdots[1,\dots,b_{i_j,r}\cdot n_j]\right);$$

for $p > \sqrt{(\max_{i,h} b_{i,h})(m+\epsilon')D}$, this equals $\#\{(j,h) : p \le b_{i_j,h}n_j\}$.

Therefore, by the prime number theorem and the definition of $\tau^{\flat\flat}(\mathbf{b})$, we have

$$\lim_{d\to\infty}\lim_{D\to\infty}\left\{ \frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)}\sum_p v_{p,\mathbf{n}}^{\flat}\log p}{dm^dD^{d+1}} \right\} \le \frac{1}{2}(m+\epsilon')\tau^{\flat\flat}(\mathbf{b}) + o_{\epsilon''\to 0}(1). \tag{8.2.33}$$

Combining (8.2.32) and (8.2.33), we get (noting that $\epsilon'' \to 0$ has, by implication, $\epsilon' \to 0$):

$$\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d}\sum_p v_{p,\mathbf{n}}^{\flat}\log p}{dm^d D^{d+1}}\right\} \le \frac{1}{2}m\tau^{\flat\flat}(\mathbf{b}). \tag{8.2.34}$$

Finally, we turn to $v_{p,\mathbf{n}}^{\sharp}$. For any $\mathbf{n} \in \mathcal{V}_D^d \subset [0,(m+\epsilon')D]^d$, at a given p, we defined $v_{p,\mathbf{n}}^{\sharp}$ as the maximal valuation of the denominators of the coefficients of all monomials $\prod_{j=1}^d x_j^{m_j}/m_j^{e_{i_j}}$ ranging over all $\mathbf{i} \in V_m^d(\epsilon)$ and all exponent vectors $\mathbf{m}$ such that $(n_j - D)^+ \le m_j \le n_j$ for all $1 \le j \le d$. It satisfies

$$\begin{aligned} v_{p,\mathbf{n}}^{\sharp} &\le \max_{\mathbf{i}\in V_m^d(\epsilon)}\left\{\sum_{j=1}^d e_{i_j}\operatorname{val}_p([\max\{n_j-D,1\},\dots,n_j])\right\} \\ &\le \left(\max_{1\le i\le m} e_i\right)\sum_{j=1}^d \operatorname{val}_p\left([\max\{n_j-D,1\},\dots,n_j]\right). \end{aligned} \tag{8.2.35}$$

It is here that we use the cutoff parameter ξ of the defining formula (6.0.5). Summing over all $p \ge \xi D$, we derive the estimate

$$\sum_{p\ge\xi D} v_{p,\mathbf{n}}^{\sharp}\log p \le \left(\max_{1\le i\le m} e_i\right)\sum_{j=1}^d\sum_{p\ge\xi D}\operatorname{val}_p([\max\{n_j-D,1\},\dots,n_j])\log p. \tag{8.2.36}$$

Since the $\mathbf{n} \in P_d^{\epsilon''}((m+\epsilon')D)$ have uniformly distributed components up to normalized discrepancy $\le \epsilon''$, Lemma 5.0.4 with the prime number theorem yields

$$\begin{aligned} &\sum_{\mathbf{n}\in\mathcal{V}_D^d\cap P_d^{\epsilon''}((m+\epsilon')D)}\left\{\sum_{p\ge\xi D} v_{p,\mathbf{n}}^{\sharp}\log p\right\} \\ &\le d\left(\max_{1\le i\le m} e_i\right)(m+\epsilon')^{d-1}D^{d+1}I_{\xi}^{m+\epsilon'}(\xi)\left(1+O(\sqrt{\epsilon''})+o(D^{d+1})\right). \end{aligned} \tag{8.2.37}$$

The complementary meagre set of $\mathbf{n}$ is once again handled by the trivial estimate:

$$\begin{aligned} \sum_{\mathbf{n}\in\mathcal{V}_D^d\cap B_d^{\epsilon''}((m+\epsilon')D)}\left\{\frac{\sum_{p\ge\xi D} v_{p,\mathbf{n}}^{\sharp}\log p}{dm^dD^{d+1}}\right\} &= O\left(\left(\max_{1\le i\le m} e_i\right)(m+\epsilon')e^{-c(\epsilon'')d}\right) \\ &= o_{d\to\infty}(1). \end{aligned} \tag{8.2.38}$$

Therefore, taking $d \to \infty$ and noting again that $\epsilon'' \to 0$ entails $\epsilon' \to 0$, we arrive at the limit majorization

$$\lim_{d\to\infty}\lim_{D\to\infty}\left\{\frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d}\sum_{p\ge\xi D} v_{p,\mathbf{n}}^{\sharp}\log p}{dm^dD^{d+1}}\right\} \le \frac{(\max_{1\le i\le m} e_i)\, I_{\xi}^m(\xi)}{m}. \tag{8.2.39}$$

It remains to estimate the $p \le \xi D$ contribution. Here we use the fact that the multi-index $\mathbf{i} \in V_m^d(\epsilon)$ is ϵ-balanced. Therefore, once again by the prime number theorem (indicating that we only count $\operatorname{val}_p = 1$), we have asymptotically as $D \to$

∞:

$$\begin{aligned}\sum_{p\leq \xi D} v^{\sharp}_{p,\mathbf{n}}\log p &\leq \sum_{p\leq \xi D}\left\{\max_{\mathbf{i}\in V^d_m(\epsilon)}\sum_{j=1}^{d} e_{i_j}\,\mathrm{val}_p([1,\ldots,n_j])\right\}\log p\\ &\leq \sum_{p\leq \xi D}\left\{\max_{\mathbf{i}\in V^d_m(\epsilon)}\sum_{j=1}^{d} e_{i_j}\right\}\log p + o(D)\\ &= d\xi D\left(\frac{\sum_{i=1}^m e_i}{m}+O(\epsilon)\right)+o(D).\end{aligned}$$

Therefore, once we let $D\to\infty$ followed by $d\to\infty$ and then $\epsilon''\to 0$ (entailing $\epsilon'\to 0$), we derive

$$\lim_{d\to\infty}\lim_{D\to\infty}\frac{\sum_{\mathbf{n}\in\mathcal{V}^d_D} v^{\sharp}_{p,\mathbf{n}}\log p}{dm^dD^{d+1}} \leq \frac{\xi(\sum_{i=1}^m e_i)+O(\epsilon)+(\max_{1\leq i\leq m} e_i)I^m_{\xi}(\xi)}{m}. \tag{8.2.40}$$

Conclusion of the proof of Theorem 8.0.1. We derive the desired bound from Bost's slopes inequality (7.2.14), upon collecting Lemma 8.2.5 and the estimates (8.2.28), (8.2.34), and (8.2.40), and finally by letting $\epsilon\to 0$ in the end.

The proof of (8.0.5) differs only in the archimedean evaluation height estimate, upon replacing $\overline{\mathcal{L}}$ by $\overline{\mathcal{L}}'=\prod_{h=0}^{l}\overline{\mathcal{L}}_{r_h}^{\otimes s_h}$ and inputting the bound from § 7.6. □

Remark 8.2.41. In the case (such as we have in all our applications in this paper) that the m-dimensional $\mathbf{Q}(x)$-vector space $\mathrm{Span}_{\mathbf{Q}(x)}\{f_1,\ldots,f_m\}$ is closed under differentiation, we can apply the more elementary Shidlovsky lemma (Theorem 3.2.8, which is much easier to prove and effectivize than Theorem 3.2.13); in this situation, the statement of Lemma 3.2.14 applies even with $\epsilon=0$. In this situation, the large deviations bound quoted from Theorem 4.2.1 in the proof can be replaced by the most rudimentary weak law of large numbers. △

Remark 8.2.42. The proof immediately gives the following formal generalization to a number field K. For each $\sigma: K\hookrightarrow\mathbf{C}$, we consider a holomorphic mapping $\varphi_\sigma:(\overline{\mathbf{D}},0)\to(\mathbf{C},0)$ with $\varphi'_\sigma(0)\neq 0$. Assume there exists an m-tuple $f_1,\ldots,f_m\in K[\![x]\!]$ of $K(x)$-linearly independent formal functions with denominator types of the form

$$f_i(x)=a_{i,0}+\sum_{n=1}^{\infty}a_{i,n}\frac{x^n}{n^{e_i}[1,\ldots,b_{i,1}\cdot n]\cdots[1,\ldots,b_{i,r}\cdot n]},\qquad a_{i,n}\in\mathcal{O}_K,$$

where $e_i, b_{i,j}$ are the same as in Theorem 6.0.2, and such that for all $i\in\{1,\ldots,m\}$ and $\sigma:K\hookrightarrow\mathbf{C}$,we have $f_i(\varphi_\sigma(z))\in\mathbf{C}[\![z]\!]$ convergent on $|z|<1$. If

$$\frac{1}{[K:\mathbf{Q}]}\sum_{\sigma:K\hookrightarrow\mathbf{C}}\log|\varphi'_\sigma(0)|>\sigma_m,$$

then all f_i are holonomic functions, and

$$m\leq\frac{\sum_{\sigma:K\hookrightarrow\mathbf{C}}\iint_{\mathbf{T}^2}\log|\varphi_\sigma(z)-\varphi_\sigma(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{(\sum_{\sigma:K\hookrightarrow\mathbf{C}}\log|\varphi'_\sigma(0)|)-[K:\mathbf{Q}](\tau^{\flat\flat}(\mathbf{b})+\tau^{\sharp}(\mathbf{e}))}.$$

The convexity improvements also extend in the obvious way. △

8.3. **The slopes method in Theorem 6.0.2.** We showed in § 8.1 that Theorem 7.0.1 formally implies the corresponding particular case Corollary 6.0.14 of Theorem 6.0.2. In this brief section, we comment how Theorem 6.0.2 can be more directly recovered in the framework of the preceding proof.

In constructing the Euclidean lattice E_D, in addition to only considering the split-variable products $f_{\mathbf{i}}$ with $\mathbf{i} \in V_m^d(\epsilon)$, we may also — as in the Thue–Siegel lemma construction in § 6.2 — constrict the monomials $\mathbf{x}^{\mathbf{k}}$ to have exponent vectors $\mathbf{k}$ with uniformly distributed components $\{k_i\}$. More precisely, define the free $\mathbf{Z}$-module:

$$E_D := \bigoplus_{\mathbf{i} \in V_m^d(\epsilon),\, \mathbf{k}/D \in P_\epsilon^d} f_{\mathbf{i}}\, \mathbf{Z}\mathbf{x}^{\mathbf{k}},$$

where P_ϵ^d was defined in (6.2.1). Then indeed we have the requisite double limit

$$\lim_{d\to\infty} \lim_{D\to\infty} \left\{ \frac{\mathrm{rk} E_D}{m^d D^d} \right\} = 1.$$

We equip E_D with the Euclidean metric that makes $\{f_{\mathbf{i}}\mathbf{x}^{\mathbf{k}}\}_{\mathbf{i}\in V_m^d(\epsilon), \mathbf{k}/D \in P_\epsilon^d}$ an orthonormal basis.

In Theorem 6.0.2, we are given a set of $l+1$ holomorphic mappings (6.0.7), and corresponding division point parameters $0 = \gamma_0 < \gamma_1 < \ldots < \gamma_l < \gamma_{l+1} := m$. Like in § 6.4, the meaning of these numbers is that we will use the Poisson–Jensen formula for $\varphi_k(z_j)$ for the unique $k = k(j)$ determined by $n_j/D \in [\gamma_k, \gamma_{k+1})$; we use $\varphi_l(z_j)$ for $n_j/D \in [m, m+\epsilon')$. We continue to use $\mathcal{V}_D^d$ to denote the set of $\mathbf{n}$ such that $\operatorname{rank} E_D^{(\mathbf{n})}/E_D^{(\mathbf{n}^+)} = 1$; recall that for $D \gg 1$, we have $\mathcal{V}_D^d \subset [0, (m+\epsilon')D]^d$. For all $\mathbf{n} \in \mathcal{V}_D^d$, the ensuing evaluation height estimate is

$$\begin{aligned} h_\infty(\psi_D^{(\mathbf{n})}) \leq D \int_{\mathbf{T}^d} \max_{\mathbf{t} \in P_\epsilon^d} \left\{ \sum_{j=1}^d t_j \log |\varphi_{k(j)}(z_j)| \right\} \mu_{\mathrm{Haar}} \\ - |\mathbf{n}| \log |\varphi_l'(0)| - \sum_{j=1}^d n_j \log |\varphi_{k(j)}'(0)/\varphi_l'(0)| + o(D). \end{aligned}$$

As in any case we have the trivial bound

$$\max_{\mathbf{t} \in P_\epsilon^d} \left\{ \sum_{j=1}^d t_j \log |\varphi_{k(j)}(z_j)| \right\} \leq d \max_{k, \mathbf{T}} \log |\varphi_k|,$$

an argument similar to § 8.2.11 shows

$$\begin{aligned} \lim_{d\to\infty} \lim_{D\to\infty} \left\{ \frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d} h_\infty(\psi_D^{(\mathbf{n})})}{d m^d D^{d+1}} \right\} &\leq \lim_{d\to\infty} \lim_{D\to\infty} \left\{ \frac{\sum_{\mathbf{n}\in\mathcal{V}_D^d \cap P_d^{\epsilon''}((m+\epsilon')D)} h_\infty(\psi_D^{(\mathbf{n})})}{d m^d D^{d+1}} \right\} \\ &\leq \lim_{d\to\infty} \left\{ \frac{1}{d} \int_{\mathbf{T}^d} \max_{\substack{\mathbf{t}\in P_\epsilon^d, \\ \mathbf{n} \in P_d^{\epsilon''}((m+\epsilon')D)}} \left\{ \sum_{j=1}^d t_j \log |\varphi_{k(j)}(z_j)| \right\} \mu_{\mathrm{Haar}} \right\} \\ &\quad - \frac{m}{2} \log |\varphi_l'(0)| - \frac{1}{2} \sum_{k=0}^{l} (\gamma_{k+1}^2 - \gamma_k^2) \log \frac{|\varphi_k'(0)|}{|\varphi_l'(0)|} + O(\epsilon' + \sqrt{\epsilon''}) \end{aligned}$$

Given $\mathbf{n} \in P_d^{\epsilon''}((m+\epsilon')D)$, asymptotically as $d \to \infty$, almost all $\mathbf{z} \in \mathbf{T}^d$ have the property that for every $k \in \{0,1,\dots,l\}$, the set $\{z_j\}_{k(j)=k}$ is equidistributed in the uniform measure μ_{Haar} of $\mathbf{T}$. Mirroring § 6.4, we thus define a function $\Phi_{\boldsymbol{\varphi},\boldsymbol{\gamma}}$ on $\mathbf{T}$ by the piecewise splicing rule[35]

$$\Phi(e^{2\pi i t})_{\boldsymbol{\varphi},\boldsymbol{\gamma}} := \varphi_k\left(e^{2\pi i \frac{mt-\gamma_k}{\gamma_{k+1}-\gamma_k}}\right), \text{ for } t \in [\gamma_k/m, \gamma_{k+1}/m).$$

In (6.0.8), we have $g_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t) = \log|\Phi_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(e^{2\pi i t})|$. Then, as in § 6.5.15, we have

$$\lim_{\epsilon\to 0}\lim_{\epsilon',\epsilon''\to 0}\lim_{d\to\infty}\left\{\frac{1}{d}\int_{\mathbf{T}^d}\max_{\substack{\mathbf{t}\in P_\epsilon^d,\\ \mathbf{n}\in P_d^{\epsilon''}\left((m+\epsilon')D\right)}}\left\{\sum_{j=1}^d t_j \log|\varphi_{k(j)}(z_j)|\right\}\mu_{\mathrm{Haar}}\right\}$$
$$= \int_0^1 t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt.$$

The argument for the non-archimedean estimate is the same as § 8.2.29, and we recover the thesis of Theorem 6.0.2:

$$m \le \frac{\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt + \frac{1}{m}\sum_{k=1}^l\left\{\gamma_k^2\log\frac{|\varphi_k'(0)|}{|\varphi_{k-1}'(0)|}\right\}}{\log|\varphi_l'(0)| - (\tau^{\flat\flat}(\mathbf{b}) + \tau^{\sharp}(\mathbf{e}))}. \tag{8.3.1}$$

Let us for concreteness now specialize to the setup $\varphi_k(z) := \varphi(r_k z)$ of § 7.4; the argument there can equally be adapted to the general situation. If now we select our division parameters $\boldsymbol{\gamma}$ to be the slopes $\gamma_k := \beta_k(s_h^*)$ in Theorem 7.6.4 (assume the linear algebra condition of that theorem to be satisfied), then by § 8.1, we have

$$\begin{aligned}
\int_0^1 2t\cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\,dt &+ \frac{1}{m}\sum_{k=1}^l \gamma_k^2\log\frac{r_k}{r_{k-1}}\\
&= \int_0^1 2t(\log|\Phi|)^*(e^{2\pi i t})\,dt - \frac{1}{m}\sum_{k=0}^l(\gamma_{k+1}^2-\gamma_k^2)\log r_k\\
&\ge \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' - \frac{1}{m}\sum_{k=0}^l(\beta_{k+1}^2-\beta_k^2)\log r_k\\
&= \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' + \frac{1}{m}\sum_{k=1}^l\beta_k^2(\log r_k - \log r_{k-1})\\
&= \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' + \frac{1}{m}\sum_{k=1}^l\beta_k(\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}' - \overline{\mathcal{L}}_{r_{k-1}}\cdot\overline{\mathcal{L}}')\\
&= \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}' + \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1 - \frac{1}{m}\sum_{k=0}^l(\beta_{k+1}-\beta_k)\overline{\mathcal{L}}_{r_k}\cdot\overline{\mathcal{L}}'\\
&= \overline{\mathcal{L}}'\cdot\overline{\mathcal{L}}_1,
\end{aligned}$$

which is the numerator of the bound that we obtained in § 8.2. In practice, § 8.1.16 suggests these two bounds to be pretty close. In particular, numerically speaking,

[35]Note that this function is different from the multivariable Φ in § 6.4.

we do not need γ_k or β_k to be exact the heuristically optimal choice. The bound (8.3.1) holds for any choice of $\boldsymbol{\gamma} = \{\gamma_k\}$.

9. The relationship between $Y(2)$ and $Y_0(2)$

There is a natural identification of $Y(2)$ with $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ given by the coordinate λ with equation (1.2.8). If we let $Y_0(2)$ denote the modular curve of level $\Gamma_0(2)$, then $Y_0(2)$ is also rational with a hauptmodul

$$h := \lambda + \frac{\lambda}{\lambda - 1} = -256q \prod_{n=1}^{\infty} (1 + q^n)^{24} = -256 \cdot \frac{\Delta(2\tau)}{\Delta(\tau)}, \quad q = e^{2\pi i \tau}, \tag{9.0.1}$$

where this time we write $q = e^{2\pi i\tau}$ in comparison to $q = e^{\pi i\tau}$ in equation (1.2.8). Just as with λ, we also view h by abuse of notation as a function of the parameter $q \in \mathbf{D}$. The parameter h gives an identification of $Y_0(2)$ with $\mathbf{P}^1 \smallsetminus \{0, \infty\}$. The map $Y(2) \to Y_0(2)$ of modular curves is smooth as a map of algebraic stacks, but not of the underlying coarse moduli spaces. To properly account for this, it is better to remember that $Y_0(2)$ has an elliptic point of order 2 at $h = 4$, which is the branch point of the double covering $\lambda \mapsto h = \lambda^2/(\lambda - 1)$, with $\lambda = 2$ for its unique preimage: the ramification divisor of the branched covering $Y(2) \to Y_0(2)$ as algebraic curves. On the stacks level, $Y(2) \to Y_0(2)$ is an étale map which is a Galois covering of degree 2, and so there is a natural relation between invariant functions on $Y(2)$ and functions on $Y_0(2)$.

However, as we shall see below, this relationship also respects some arithmetic properties of the corresponding power series expansions. First we remark that the transformation

$$w : x \mapsto \frac{x}{x-1} \in \mathbf{Z}[\![x]\!] \tag{9.0.2}$$

is an involution of $\mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ that preserves 0 and swaps 1 and ∞. The integrality properties of this map (and its inverse) means that if $f(x) \in \mathbf{Q}[\![x]\!]$ has a certain denominator type then so does $f(w(x))$. Second, we note that the map w of equation (9.0.2) is precisely the non-trivial Galois automorphism of (the function field of) $Y(2)$ over $Y_0(2)$.

Lemma 9.0.3. *Let $S \subset \mathbf{P}^1 \smallsetminus \{0, 1, \infty\}$ be a finite set invariant under the involution w of equation (9.0.2), and define $T \subset \mathbf{P}^1 \smallsetminus \{0, \infty\}$ to be the image of S under the map $x \mapsto y$, where*

$$y := x + w(x) = x + \frac{x}{x-1} = \frac{x^2}{x-1}. \tag{9.0.4}$$

Consider $c_1, \ldots, c_r \in [0, \infty)$, and let $f(x) \in \mathbf{Q}[\![x]\!]$ be a power series of the form

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{\prod_{i=1}^r [1, \ldots, c_i n]} \in \mathbf{Q}[\![x]\!], \qquad a_n \in \mathbf{Z} \quad \forall n \in \mathbf{N}, \tag{9.0.5}$$

which converges on a neighborhood of $x = 0$ and continues analytically as a holomorphic function along all paths in $\mathbf{P}^1 \smallsetminus \{0, 1, S, \infty\}$. Then:

1. *The function $f(w(x)) \in \mathbf{Q}[\![x]\!]$ is also of the form (9.0.5).*
2. *If $f(x) = f(w(x))$, then we may use (9.0.4) to formally write $f(x)$ as a power series $f(x) = F(y) \in \mathbf{Q}[\![y]\!]$ that satisfies*

$$2F(y) = \sum_{n=0}^{\infty} b_n \frac{y^n}{\prod_{i=1}^r [1, \ldots, 2c_i n]} \in \mathbf{Q}[\![y]\!], \qquad b_n \in \mathbf{Z} \quad \forall n \in \mathbf{N}, \tag{9.0.6}$$

converges in a neighborhood of $y = 0$, continues analytically as a holomorphic function along all paths in $\mathbf{P}^1 \smallsetminus \{0,4,\infty,T\}$, and has finite local monodromy[36] of order dividing 2 around the point $y = 4$.

Conversely, if $2F(y)$ has the form (9.0.6), then $2F\left(x+\dfrac{x}{x-1}\right)$ has the form (9.0.5), and if $F(y)$ has the analyticity properties on $\mathbf{P}^1 \smallsetminus \{0,4,\infty,T\}$ spelled out in (2), *then $F\left(x+\dfrac{x}{x-1}\right)$ has the analytic continuation property on $\mathbf{P}^1 \smallsetminus \{0,1,S,\infty\}$.*

Proof. The function-theoretic claims follow directly from Galois theory and the fact that $x \mapsto w(x)$ is the automorphism of $Y(2)$ over $Y_0(2)$. Hence it suffices to establish the claims concerning integrality. Property (1) is clear from the integral coefficients in the expansion $w(x)^n = (-1)^n x^n (1-x)^{-n} \in x^n \mathbf{Z}[\![x]\!]$ together with the remark that the denominator type $\prod_{i=1}^r [1,\ldots,c_i n]$ is nested by division under $n \mapsto n+1$. Let x and y be related by the identity (9.0.4). Since both the elementary symmetric functions in x and $w(x) = x/(x-1)$ are equal to $y = x + w(x) = xw(x)$, we may define polynomials $P_n(y)$ by the rule

$$P_n(y) := x^n + (x/(x-1))^n. \tag{9.0.7}$$

Then $P_0(y) = 2,\ P_1(y) = y$, and there is the elementary recurrence

$$P_n(y) = yP_{n-1}(y) - yP_{n-2}(y).$$

We find that $P_n(y)$ has degree n and vanishes at $y = 0$ to order $\lceil n/2 \rceil$. Let us suppose now that we have a function $f(x) = \sum A_n x^n$ whose coefficients A_n are rational numbers with $a_n := A_n \prod_{i=1}^r [1,\ldots,c_i n] \in \mathbf{Z}$. Then, in property (2) under proof, we exploit the assumption $f(w(x)) = f(x)$ to write

$$f(x) + f\left(\frac{x}{x-1}\right) = 2F(y) = \sum A_k P_k(y) =: \sum B_n y^n.$$

The middle equality defines a legitimate $\mathbf{Q}[\![y]\!]$ series since $P_k(y)$ is divisible by $y^{\lceil k/2 \rceil}$, and it can be taken as a definition. To be more precise, all the nonzero coefficients of the polynomial $P_k(y)$ occur in the degree range $[k/2, k]$, and they are integers. Thus $P_k(y)$ contributes to the y^n term only for $k \in [n, 2n]$, and $b_n := B_n \prod_{i=1}^r [1,2,\ldots,2c_i n] \in \mathbf{Z}$. Conversely, since $x + \dfrac{x}{x-1} = \dfrac{x^2}{x-1}$, if we write

$$\sum A_n x^n = \sum B_k \left(x + \frac{x}{x-1}\right)^k = \sum B_k \frac{x^{2k}}{(x-1)^k},$$

then the terms on the right-hand side contributing to A_n occur only for $k \le n/2$. $\square$

Motivated by this lemma, we have the following:

---

[36] In this generality, by a "finite local monodromy of order dividing 2" we simply mean that if $\gamma : (0,1] \to \mathbf{P}^1 \smallsetminus \{0,4,\infty,T\}$ is any path with origin $\lim_{t\to 0^+} \gamma(t) = 0$, and π is a simple loop around $y = 4$ in $\mathbf{P}^1 \smallsetminus \{0,4,\infty,T\}$ based at the endpoint $\gamma(1)$, then the analytic continuations of $f(y)$ at the ends of the concatenated paths γ and $\pi^2 \cdot \gamma$ are equal. This, of course, agrees with the usual notion in the special (finite-dimensional) case of a local system on $\mathbf{P}^1 \smallsetminus \{0,4,\infty,T\}$.

Definition 9.0.8. Let $F(x) \in \mathbf{Q}[\![x]\!]$. We define the plus and minus *symmetrization* functions $F^+(y)$ and $F^-(y)$ to be the elements of $\mathbf{Q}[\![y]\!]$ such that

$$\begin{aligned} F^+(y) &= F(x) + F\left(\frac{x}{x-1}\right), \\ F^-(y) &= \left(x - \frac{x}{x-1}\right)\left(F(x) - F\left(\frac{x}{x-1}\right)\right), \end{aligned} \tag{9.0.9}$$

where $y := x + \dfrac{x}{x-1} = \dfrac{x^2}{x-1}$. △

We connect these symmetrizations to the analytic resolvents $\varphi^* f \in \mathcal{O}(\mathbf{D})$ in the context of the arithmetic holonomy bounds. We firstly introduce an *ad hoc* definition (which will only be used in Lemma 9.0.13):

Definition 9.0.10. A *holonomic descent datum* is a tuple

$$\mathcal{R}_f = \left(U_{Y(2)}, \Sigma^0_{Y(2)}, \Sigma^1_{Y(2)}, f\right)$$

consisting of:

(1) A contractible open neighborhood $0 \in U_{Y(2)} \subset \mathbf{C} \smallsetminus \{2\}$ which is invariant under the involution w.
(2) Finite subsets $\Sigma^0_{Y(2)} \subset U_{Y(2)}$ and $\Sigma^1_{Y(2)} \subset Y(2) = \mathbf{C} \smallsetminus \{0, 1\}$, both invariant under the involution w.
(3) A holomorphic function $f \in \mathcal{O}\left(U_{Y(2)}\right)$ which is w-invariant ($f(w(x)) = f(x)$) and analytically continuable as a holomorphic function along all paths in $\mathbf{P}^1 \smallsetminus \{0, 1, \Sigma^0_{Y(2)}, \Sigma^1_{Y(2)}, \infty\}$.

To every such datum $\mathcal{R}_f$, we attach a *quotient datum*

$$\mathcal{Q}_F = \left(U_{Y_0(2)}, \Sigma^0_{Y_0(2)}, \Sigma^1_{Y_0(2)}, F\right)$$

as follows. By expressing the w-invariant power series expansion $f(x) \in \mathbf{C}[\![x]\!]$ formally into $y := x + w(x)$, we have attached as[37] in Lemma 9.0.3 a unique formal power series $F(y) = F_f(y) \in \mathbf{C}[\![y]\!]$ such that

$$F(y) = F(x + w(x)) = F\left(x + \frac{x}{x-1}\right) = f(x). \tag{9.0.11}$$

In a similar manner, we define $\Sigma^0_{Y_0(2)}$, $\Sigma^1_{Y_0(2)}$, and $U_{Y_0(2)}$ to be the images of $\Sigma^0_{Y(2)}$, $\Sigma^1_{Y(2)}$, and $U_{Y(2)}$, under the map

$$x \mapsto y := x + w(x) = x + \frac{x}{x-1}$$

of (9.0.4). △

For brevity we also adopt the following definition, which formalizes the idea of the univalent leaves from Proposition 2.9.3.

Definition 9.0.12. Consider two pointed Riemann surfaces (D, O) and (X, P) and an open neighborhood $P \in U \subset X$. A holomorphic mapping $\varphi : (D, O) \to (X, P)$

[37] Except that, now in the analytic context, we can assume $f \in \mathbf{C}[\![x]\!]$ rather than $f \in \mathbf{Q}[\![x]\!]$; the proof, of course, is the same.

has a univalent leaf over U at O if φ maps the connected component of $\varphi^{-1}(U)$ containing O conformally isomorphically onto U:

$$\varphi \,:\, \left(\varphi^{-1}(U)\right)_O \xrightarrow{\simeq} U.$$

We refer to $\left(\varphi^{-1}(U)\right)_O \subset D$ itself as the univalent leaf (at O over U). △

Lemma 9.0.13. *Let $\mathcal{R}_f$ be a holonomic descent datum with quotient $\mathcal{Q}_F$. Let $\varphi_{Y(2)}$ be a holomorphic mapping that obeys*

$$\varphi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \{1, \Sigma^1_{Y(2)}\}, \quad \varphi^{-1}_{Y(2)}(0) = \{0\},$$

and which has a univalent leaf $\left(\varphi^{-1}_{Y(2)}\left(U_{Y(2)}\right)\right)_0$ over $U_{Y(2)}$ at $0 \in \mathbf{D}$ containing all the pre-images of $\Sigma^0_{Y(2)}$ under $\varphi_{Y(2)}$.

Suppose that

$$w(\varphi_{Y(2)}(z)) = \varphi_{Y(2)}(-z). \tag{9.0.14}$$

The pullback $\varphi^*_{Y(2)}f$ is holomorphic on $\overline{\mathbf{D}}$. If $\varphi_{Y_0(2)}$ is the holomorphic map

$$\varphi_{Y_0(2)}(z) := \varphi_{Y(2)}\left(\sqrt{z}\right) + \varphi_{Y(2)}\left(-\sqrt{z}\right) \in \mathcal{O}(\overline{\mathbf{D}}), \tag{9.0.15}$$

then:

(1) *$\varphi^{-1}_{Y_0(2)}(0) = \{0\}$.*
(2) *The range of $\varphi_{Y_0(2)}$ omits $\Sigma^1_{Y_0(2)}$.*
(3) *The ramification indices of $\varphi_{Y_0(2)}$ are even at all points of the fiber $\varphi^{-1}_{Y_0(2)}(4)$.*
(4) *The neighborhood $U_{Y_0(2)} \ni 0$ is a contractible domain, and $\varphi_{Y_0(2)}$ has a univalent leaf over $U_{Y_0(2)}$ at $0 \in \mathbf{D}$, which furthermore contains all the pre-images of $\Sigma_{Y_0(2)}$ under $\varphi_{Y_0(2)}$.*
(5) *$F|_{U_{Y_0(2)}} \in \mathcal{O}\left(U_{Y_0(2)}\right)$ is holomorphic, and the following relation holds:*

$$\begin{aligned} F\left(\varphi_{Y_0(2)}(z)\right) &= f\left(\varphi_{Y(2)}(\sqrt{z})\right) = f\left(w\left(\varphi_{Y(2)}(\sqrt{z})\right)\right) \\ &= \frac{f\left(\varphi_{Y(2)}(\sqrt{z})\right) + f\left(\left(\varphi_{Y(2)}(-\sqrt{z})\right)\right)}{2} \in \mathcal{O}(\overline{\mathbf{D}}). \end{aligned} \tag{9.0.16}$$

In particular, the pullback of F by $\varphi_{Y_0(2)}$ is holomorphic on $\overline{\mathbf{D}}$.

Conversely, every holomorphic mapping $\varphi_{Y_0(2)} \in \mathcal{O}(\overline{\mathbf{D}})$ obeying the conditions (1) *through* (4) *determines through* (9.0.15) *a unique pair $\left\{\varphi_{Y(2)}, w \circ \varphi_{Y(2)}\right\}$ of holomorphic mappings*

$$\varphi : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \{1, \Sigma^1_{Y(2)}\}, \quad \varphi^{-1}(0) = \{0\}$$

subject to $w\left(\varphi(z)\right) = \varphi(-z)$.

Proof. The holomorphy of $\varphi^*_{Y(2)}f$ on $\overline{\mathbf{D}}$ follows directly from Proposition 2.9.3 with $\Omega := \left(\varphi^{-1}_{Y(2)}(U_{Y(2)})\right)_0$, as $f \in \mathcal{O}\left(U_{Y(2)}\right)$. We observe that $U_{Y_0(2)}$ — a domain, by the open mapping theorem — is also a topological disc, as $U_{Y(2)} \subset \mathbf{C} \smallsetminus \{2\}$ while the map $y := x^2/(x-1)$ has $x \in \{0, 2\}$ for its only ramification points, with branching values $y \in \{0, 4\}$.

Assume now the symmetries $f(x) = f(w(x))$ and $w(\varphi_{Y(2)}(z)) = \varphi_{Y(2)}(-z)$, and define the manifestly holomorphic map $\varphi_{Y_0(2)} \in \mathcal{O}(\overline{\mathbf{D}})$ by (9.0.15). It is the

$w(\varphi_{Y(2)}(z)) = \varphi_{Y(2)}(-z)$ symmetry that allows to descend the analytic data to the $Y_0(2)$ picture, as the plus-symmetrization of $\varphi_{Y(2)}$:

$$\begin{aligned}\varphi_{Y_0(2)}(z) &:= \varphi_{Y(2)}\left(\sqrt{z}\right) + \varphi_{Y(2)}\left(-\sqrt{z}\right) \in \mathcal{O}(\overline{\mathbf{D}}) \\ &= \varphi_{Y(2)}\left(\sqrt{z}\right) + w\left(\varphi_{Y(2)}\left(\sqrt{z}\right)\right) \\ &= \frac{\varphi_{Y(2)}(\sqrt{z})^2}{\varphi_{Y(2)}(\sqrt{z}) - 1}.\end{aligned} \tag{9.0.17}$$

The second line — together with the definitional fact that $\Sigma^1_{Y(2)}$ is the full inverse image of $\Sigma^1_{Y_0(2)}$ under the double covering map $y = x + w(x)$ — shows that the range of $\varphi_{Y_0(2)}$ omits the set $\Sigma^1_{Y_0(2)}$, as $\varphi_{Y(2)}$ omits the corresponding set $\Sigma^1_{Y(2)}$. The third line shows that $\varphi_{Y_0(2)}$ satisfies $\varphi^{-1}(0) = \{0\}$ and has even ramification indices at all points in the fiber $\varphi^{-1}_{Y_0(2)}(4) = \varphi^{-1}_{Y(2)}(2)$. Applying $F(x + w(x)) = f(x) = f(w(x))$ for the $x = \varphi_{Y(2)}(\sqrt{z})$ given on the second line in (9.0.17), we get (9.0.16), and in particular the holomorphy of $\varphi^*_{Y_0(2)}F \in \mathcal{O}(\overline{\mathbf{D}})$. Lastly, the holomorphy $F|_{U_{Y_0(2)}} \in \mathcal{O}\left(U_{Y_0(2)}\right)$ follows directly from the corresponding holomorphy $f|_{U_{Y(2)}} \in \mathcal{O}\left(U_{Y(2)}\right)$ thanks to the defining equation $F(x + w(x)) = f(x)$ and the definition of $U_{Y_0(2)}$ as the $x + w(x)$ image of $U_{Y(2)}$.

For the converse, we get by the formal binomial expansion — choosing any branch for the square root signs — a power series $\varphi_{Y(2)} \in \mathbf{C}[\![z]\!]$ from resolving the quadratic relation on the third line in (9.0.17) with z changed to z^2:

$$\varphi_{Y(2)}(z) := \frac{\varphi_{Y_0(2)}(z^2) + \sqrt{\varphi_{Y_0(2)}(z^2)} \cdot \sqrt{\varphi_{Y_0(2)}(z^2) - 4}}{2}. \tag{9.0.18}$$

The conditions on $\varphi^{-1}_{Y_0(2)}(0) = \{0\}$ and on even ramification indices for $\varphi_{Y_0}(2)$ along $\varphi^{-1}_{Y_0(2)}(4)$ show that the formal function (9.0.18) is in fact holomorphic on a neighborhood of $\overline{\mathbf{D}}$, and satisfies $\varphi^{-1}_{Y(2)}(0) = \{0\}$ and $w\left(\varphi_{Y(2)}(z)\right) = \varphi_{Y(2)}(-z)$. The other choice of the square roots sign in (9.0.18) leads to the argument sign swap $\varphi_{Y(2)}(-z)$, and the pair $\left\{\varphi_{Y(2)}, w \circ \varphi_{Y(2)}\right\}$ is uniquely determined from $\varphi_{Y_0(2)}$ and satisfies (9.0.17), whence the range property $\varphi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \{1, \Sigma^1_{Y(2)}\}$ is also inherited. □

We spell out as a separate corollary the case that we will use of analytic pullbacks of the hauptmodul map (9.0.1). This should be regarded as a stacky version for $Y_0(2)$ of Proposition 2.9.3 on overconvergence.

Corollary 9.0.19. *Consider an arbitrary power series $F \in \mathbf{C}[\![y]\!]$ that defines a holomorphic function on a contractible open neighborhood $0 \in U_{Y_0(2)} \subset \mathbf{C} \smallsetminus \{4\}$. Suppose $\Sigma^0_{Y_0(2)} \subset U_{Y_0(2)}$ and $\Sigma^1_{Y_0(2)} \subset \mathbf{C}$ are finite subsets such that $F(y)$ continues analytically as a holomorphic function along all paths in $y \in \mathbf{P}^1 \smallsetminus \{0, 4, \Sigma^0_{Y_0(2)}, \Sigma^1_{Y_{0(2)}}\}$ and has around $y = 4$ a finite local monodromy of order dividing 2. Let $h : \mathbf{D} \to \mathbf{C}$ be the map (9.0.1).*

Then, under any holomorphic mapping $\varphi_{Y_0(2)} : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \Sigma^1_{Y_0(2)}$ that has a univalent leaf over $U_{Y_0(2)}$ at $0 \in \mathbf{D}$ containing $\varphi^{-1}\left(\Sigma^0_{Y_0(2)}\right)$, and which factors as a composition $\varphi_{Y_0(2)} = h \circ \psi_{Y_0(2)}$ for some holomorphic $\psi_{Y_0(2)} : \overline{\mathbf{D}} \to \mathbf{D}$ with $\psi^{-1}_{Y_0(2)}(0) = \{0\}$, the pullback of F is holomorphic: $\varphi^*_{Y_0(2)}F \in \mathcal{O}(\overline{\mathbf{D}})$.

Proof. Define $U_{Y(2)} := y^{-1}\left(U_{Y_0(2)}\right)$ as the full inverse image under the map $y := x + w(x) = x^2/(x-1)$. Since $4 \notin U_{Y_0(2)}$, this neighborhood $U_{Y(2)} \ni 0$ is also contractible. Setting also $f(x) := F(y) = F(x+w(x))$ and $\Sigma^0_{Y(2)} := y^{-1}\left(\Sigma^0_{Y(2)}\right)$, $\Sigma^1_{Y(2)} := y^{-1}\left(\Sigma^1_{Y(2)}\right)$, we have thus constructed a holonomic descent datum $\mathcal{R}_f$ with quotient $\mathcal{Q}_F = \left(U_{Y_0(2)}, \Sigma^0_{Y_0(2)}, \Sigma^1_{Y_0(2)}, F\right)$.

Since (with our assumptions on $\psi_{Y_0(2)}$) the maps of the form $\varphi_{Y_0(2)} = h \circ \psi_{\psi_{Y_0(2)}}$ satisfy the conditions (1) through (4) in Lemma 9.0.13, the converse direction of the lemma then constructs a holomorphic mapping $\varphi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{C} \smallsetminus \{1, \Sigma^1_{Y(2)}\}$ with $\varphi^{-1}_{Y(2)}(0) = \{0\}$ and $w\left(\varphi_{Y(2)}(z)\right) = \varphi_{Y(2)}(-z)$, and inducing a conformal isomorphism $\varphi^{-1}_{Y(2)}\left(U_{Y(2)}\right)_0 \xrightarrow{\simeq} U_{Y(2)}$: a univalent leaf over $U_{Y(2)}$ at $0 \in \mathbf{D}$. The forward direction of Lemma 9.0.13 now proves the holomorphy $\varphi^*_{Y(2)} f \in \mathcal{O}(\overline{\mathbf{D}})$ together with the symmetrization relation (9.0.16), which in particular manifests the holomorphy $\varphi^*_{Y_0(2)} F \in \mathcal{O}(\overline{\mathbf{D}})$. □

Basic Remark 9.0.20. We combine and interpret Lemmas 9.0.3 and 9.0.13 as follows. For the remainder of our paper, we will consistently reserve the letter y to denote the covering $y := x^2/(x-1)$. Suppose given a $\mathbf{Q}(x)$-vector space $\mathcal{H}$ generated by $\mathbf{Q}[\![x]\!]$ power series of the arithmetic type (9.0.5), holomorphic on some neighborhood $U_{Y(2)} \ni 0$, and analytically continuing as holomorphic functions along all paths in $\mathbf{P}^1 \smallsetminus \{0,1,S,\infty\}$. Lemma 9.0.3 then constructs a corresponding $\mathbf{Q}(y)$-vector space $\mathcal{H}^{w=1}$ over $\mathbf{Q}(y)$ of functions on $\mathbf{P}^1 \smallsetminus \{0,4,T,\infty\}$, with at most $\mathbf{Z}/2$ local monodromy around $y = 4$, and satisfying the arithmetic condition (9.0.6). Moreover, from basic Galois theory, we have

$$\dim_{\mathbf{Q}(x)} \mathcal{H} = \dim_{\mathbf{Q}(y)} \mathcal{H}^{w=1}. \tag{9.0.21}$$

Explicitly, we have $\dim_{\mathbf{Q}(y)}(\mathbf{Q}(x)) = 2$ with a basis given by 1 and $y^- = x - \dfrac{x}{x-1}$. There is an isomorphism of $\mathbf{Q}(y)$-vector spaces $\mathcal{H} = \mathcal{H}^{w=1} \oplus \mathcal{H}^{w=-1}$ given by

$$F(x) \mapsto \left(F^+(y), F^-(y)/y^-\right) = \left(F(x) + F\left(\frac{x}{x-1}\right), F(x) - F\left(\frac{x}{x-1}\right)\right),$$

and an isomorphism $\mathcal{H}^{w=1} \to \mathcal{H}^{w=-1}$ given by multiplication by y^-. Hence

$$\dim_{\mathbf{Q}(y)} \mathcal{H}^{w=1} = \frac{1}{2}\dim_{\mathbf{Q}(y)} \mathcal{H} = \dim_{\mathbf{Q}(x)} \mathcal{H}.$$

One can now ask what happens (for example) to a holonomy bound of the form:

$$\dim_{\mathbf{Q}(x)} \mathcal{H} \leq \frac{\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|\varphi'(0)| - \sigma}, \tag{9.0.22}$$

when translated from the $Y(2)$ or $\mathbf{Q}(x)$ domain into the $Y_0(2)$ or $\mathbf{Q}(y)$ domain?

The answer to this question is that the corresponding bounds (9.0.22) are, like the dimensions (9.0.21) themselves, also equivalent in the framework of Theorem 2.5.1. Firstly, we need to make precise what we mean by the $Y(2)$ versus the $Y_0(2)$ domain in the context of formal-analytic arithmetic surfaces and arithmetic holonomy bounds. This is the content and purpose of Lemma 9.0.13. Coming from the setting of § 2.9.5, "using the $Y(2)$ domain" refers to the holomorphic mappings

$\varphi = \varphi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{C} \setminus \{1\} = Y(2) \cup \{0\}$ with $\varphi^{-1}(0) = \{0\}$, and therefore factorizing as

$$\varphi_{Y(2)} = \lambda \circ \psi_{Y(2)},$$

where $\psi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{D}$ is a holomorphic map still having $\psi_{Y(2)}^{-1}(0) = \{0\}$. The proof of this factorization (cf. [Car54, §§ 4.11, 4.12] for the details) reduces to the fact that $\tau \mapsto \lambda(\tau)$, $\tau : \mathbf{H} \to \mathbf{P}^1 \setminus \{0, 1, \infty\}$ is a universal covering map at $\tau(i) = 1/2$. On the other hand, the basic properties of the modular lambda map also include $\lambda(\tau + 1) = \lambda(\tau)/(\lambda(\tau) - 1)$ in the $\tau \in \mathbf{H}$ domain, that is $\lambda(-q) = w(\lambda(q))$ in the $q = e^{\pi i \tau} \in \mathbf{D}$ domain. Therefore, if we impose the condition $\psi_{Y(2)}(-z) = -\psi_{Y(2)}(z)$ on the map $\psi_{Y(2)}$, then the involution w acts as $w(\varphi_{Y(2)}(z)) = \varphi_{Y(2)}(-z)$. In the special case that $\psi_{Y(2)} : (\mathbf{D}, 0) \xrightarrow{\simeq} (\Psi, 0)$ is the Riemann map of a contractible domain Ψ with $0 \in \Psi \subset \overline{\Psi} \subset \mathbf{D}$, this condition simply amounts to asking for the domain Ψ to be symmetric across the origin. We further assume that the open neighborhood $U_{Y(2)}$ of the origin meets the conditions of Lemma 9.0.13: namely, $w\left(U_{Y(2)}\right) = U_{Y(2)}$, and $\varphi_{Y(2)}^{-1}\left(U_{Y(2)}\right)_0 \xrightarrow{\simeq} U_{Y(2)}$ is a univalent leaf of $\varphi_{Y(2)}$ at $0 \in \mathbf{D}$ containing all pre-images of $\Sigma_{Y(2)}^0$. We note that this property implies, but is stronger than, the corresponding property for the inner map $\psi_{Y(2)}$ in the factorization $\varphi_{Y(2)} = \lambda \circ \psi_{Y(2)}$.

We are now in the realm of Lemma 9.0.13, where we may look for a decomposition $S = \Sigma_{Y(2)}^0 \sqcup \Sigma_{Y(2)}^1$ with $\Sigma_{Y(2)}^0 \subset U_{Y(2)}$ and $\varphi_{Y(2)} : \overline{\mathbf{D}} \to \mathbf{C} \setminus \{1, \Sigma_{Y(2)}^1\}$. Under these conditions, we obtain from $\varphi_{Y(2)}$ a map $\varphi_{Y_0(2)}$ such that $\varphi_{Y(2)}^* f$ and $\varphi_{Y_0(2)}^* F$ are meromorphic on $\overline{\mathbf{D}}$ for any $f \in \mathcal{H}$ or $F \in \mathcal{H}^{w=1}$ respectively.

It is with these choices $\varphi_{Y(2)}$ and $\varphi_{Y_0(2)}$ for the analytic mapping φ, and with correspondingly the terms $\sigma := \tau(\mathbf{b})$, resp. $\sigma := \tau(2\mathbf{b}) = 2\tau(\mathbf{b})$ under the formulation of Theorem 2.5.1, that we are comparing the holonomy quotients (9.0.22) under the dictionary supplied by Lemma 9.0.3.

We now substantiate our claim that these two quotients (9.0.22) are exactly equal. The preceding analysis relies on the fact that the map $Y(2) \to Y_0(2)$ of algebraic stacks is étale. On the other hand, as the branched double covering of rational algebraic curves $X(2) \to X_0(2)$ is totally ramified over the center $h = 0$ (the cusp $\tau = i\infty$) of our formal function expansions, it follows by the projection formula in Lemma 7.4.5 that *both* the corresponding integral and conformal radius terms on the $Y_0(2)$ version of the holonomy quotient (9.0.22) are exactly scaled by the degree of that covering (which in our case is equal to two). In our basic situation, we can see this in a very direct and explicit way as follows. Let us write $H(\tau)$ for the hauptmodul h evaluated at $e^{2\pi i \tau}$, and $L(\tau)$ for λ evaluated at $e^{\pi i \tau}$, both with τ in the upper half plane $\mathbf{H}$. If $\psi_{Y(2)}(e^{i\theta}) = e^{2\pi i \tau}$ with $\tau \in \mathbf{H}$, then, making some (consistent) choice of square roots, we have

$$\varphi_{Y_0(2)}(e^{i\theta}) = h(e^{2\pi i \tau}) = H(\tau),$$

whereas

$$\varphi_{Y(2)}(e^{i\theta/2}) = \lambda(e^{\pi i \tau}) = L(\tau), \quad \varphi_{Y(2)}(-e^{i\theta/2}) = L(\tau + 1) = \frac{L(\tau)}{L(\tau) - 1}.$$

In particular, the $Y(2)$ integral involving $\log|L(\tau)-L(\sigma)|$ becomes, after dividing the integral up into four pieces, an integral of

$$\log|L(\tau)-L(\sigma)| + \log\left|\frac{L(\tau)}{L(\tau)-1} - L(\sigma)\right| + \log\left|L(\tau) - \frac{L(\sigma)}{L(\sigma)-1}\right| + \log\left|\frac{L(\tau)}{L(\tau)-1} - \frac{L(\sigma)}{L(\sigma)-1}\right|.$$

But now (using only the multiplicativity property of the logarithm and some elementary algebra) this is exactly

$$2\log\left|L(\tau) + \frac{L(\tau)}{L(\tau)-1} - L(\sigma) - \frac{L(\sigma)}{L(\sigma)-1}\right| = 2\log|H(\tau)-H(\sigma)|.$$

Taking into account the factors of 2 coming from the various scalings, this means that the integral in the $Y_0(2)$ domain is *precisely double* the integral in the $Y(2)$ domain. On the other hand, the conformal radius is also squared (this is clear for the factors of $\psi_{Y(2)}$ and $\psi_{Y'(2)}$ together with the equality $|h'(0)| = 256 = |\lambda'(0)|^2$), and so the logarithm of the conformal radius is also doubled. At the same time, in the context of Lemma 9.0.3, the invariant $\sigma = \sum_{i=1}^{r} c_i$ is *also* doubled, and so is the invariant $\tau(\mathbf{b})$ in the context of Theorem 2.5.1.

In summary, the bound (9.0.22) applied to $\dim_{\mathbf{Q}(x)}\mathcal{H}$ and $\dim_{\mathbf{Q}(y)}\mathcal{H}^{w=1}$ (which are equal by equation (9.0.21)) gives the same result in both cases. This therefore gives a (rough) equivalence between these two problems on both the arithmetic and the analytic sides. However, it is also important to note is that this crisp equivalence of the bounds only applies to the framework of the crude denominator types as stated in Theorem 2.5.1 or Lemma 9.0.3, and that there *is* still a difference once we start to consider refined denominators data, such as with the $\tau^\sharp$ from the added integrals in § 6. We shall see in § 10.3 that it can then be advantageous to perform the $\varphi_{Y(2)} \rightsquigarrow \varphi_{Y_0(2)}$ analytic descent passage that we detailed in this Basic Remark. △

10. Pure functions on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ and on $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$

The goal of this section is to write down a number of G-functions with nice integrality properties on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$, and then, using the translation discussion in § 9, on $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$ as well, where the point $y=4$ should be though of as an elliptic point of order 2. In terms of local systems on an orbifold, the proper way to think of these domains is as the modular curves $Y(2)$ in the coordinate $x=\lambda$, and respectively, $Y_0(2)$ in the coordinate $y = x^2/(x-1) = \lambda^2/(\lambda-1) = h$.

10.1. **Five functions of type $[1,\dots,n]^2$ on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$.** There are four obvious $\mathbf{Q}(x)$-linearly independent G-functions we can write down on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$

with denominator type $\tau = [1, 2, 3, \ldots, n]^2$. Namely:

$$\begin{aligned} A_1(x) &= 1, \\ A_2(x) &= -\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}, \\ A_3(x) &= \log^2(1-x) = \left(\sum_{n=1}^{\infty} \frac{x^n}{n}\right)^2, \\ A_4(x) &= \mathrm{Li}_2(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}. \end{aligned} \tag{10.1.1}$$

Clearly $A_2(x)$ additionally has type $\tau = [1, 2, 3, \ldots, n]$, and $A_3(x)$ has denominator type $[1, 2, \ldots, n][1, 2, \ldots, n/2]$. These functions are linearly independent over $\mathbf{Q}(x)$. Using symmetrizations, we also obtain 4 linearly independent functions over $\mathbf{Q}(y)$. These can be given explicitly as follows:

$$\begin{aligned} B_1(y) &= 1, \\ B_2(y) &= \sum_{n=2}^{\infty} 2y^n \cdot \frac{(n-2)!n!}{(2n)!} = 2y - 2\sqrt{y(4-y)} \arcsin\left(\frac{\sqrt{y}}{2}\right) \\ B_3(y) &= \sum_{n=1}^{\infty} y^n \cdot \frac{(n-1)!^2}{(2n)!} = 2 \arcsin(\sqrt{y}/2)^2 \\ B_4(y) &= \mathrm{Sym}^{-}\mathrm{Li}_2(y) = \left(x - \frac{x}{x-1}\right)\left(\mathrm{Li}_2(x) - \mathrm{Li}_2\left(\frac{x}{x-1}\right)\right) \\ &= -2\sqrt{y(4-y)} \int \frac{\arcsin\left(\frac{\sqrt{y}}{2}\right)}{y} \\ &= 4 \cdot \sum_{n=0}^{\infty} \frac{y^{n+1}}{16^n} \left(\sum_{k=0}^{n} \binom{2k}{k}\binom{2n-2k}{n-k} \frac{1}{(2k-1)(2n-2k+1)^2}\right) \\ &= -4y + \frac{4y^2}{9} + \frac{31y^3}{900} + \frac{389y^4}{88200} + \ldots. \end{aligned} \tag{10.1.2}$$

These functions all have denominator type subsumed by $[1, 2, \ldots, 2n]^2$ (for a more precise description, see Lemma 10.2.2 and Remark 10.2.3).

We spent a possibly embarrassing period of time believing that the four functions $A_1(x), \ldots, A_4(x)$ spanned the $\mathbf{Q}(x)$-vector space of functions on $\mathbf{P}^1 \setminus \{0, 1, \infty\}$ with denominator type $\tau = [1, 2, \ldots, n]^2$. However, there is also a fifth function one can write down. It arises more naturally in the $\mathbf{Q}(y)$-domain, namely as

$$B_5(y) = \sum_{n=1}^{\infty} y^n \cdot \frac{(n-1)!^2}{(2n-1)! \cdot (2n-1)} = y \cdot {}_3F_2\left[\begin{matrix} 1/2 & 1 & 1 \\ 3/2 & 3/2 \end{matrix}; \frac{y}{4}\right]. \tag{10.1.3}$$

The function $B_5(y)$ arises in Nesterenko's approximations [Nes16] to Catalan's constant $G = L(2, \chi_{-4})$ [Cat1882] in association with the equality

$$B_5(4) = 8G \tag{10.1.4}$$

due to Nielsen [Nie1909, page 166]. Here $B_5(y)$ is of type $[1,2,3,\ldots,2n]^2$ (and even somewhat better than this, see Lemma 10.2.2). One can easily define a corresponding function

$$A_5(x) = J(x) = x \cdot {}_3F_2\left[\begin{matrix}1/2 & 1 & 1\\ 3/2 & 3/2\end{matrix}; \frac{1}{4}\left(x + \frac{x}{x-1}\right)\right] \tag{10.1.5}$$

on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ with denominator type $\tau = [1,2,3,\ldots,n]^2$, which we originally missed! If

$$\mathcal{L} = 2x(1-x)^2\frac{d^2}{dx^2} + (2-x)(1-x)\frac{d}{dx} + 1,$$

then $\mathcal{L}J(x) = 2-2x$. We also see that

$$2x(x-1)\frac{dJ(x)}{dx} - xJ(x) = 2(1-x)\log(1-x). \tag{10.1.6}$$

From the differential equation we see that $J(x)$ is defined on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$. The solutions to the homogenous differential equation $\mathcal{L}(F) = 0$ are given by

$$A(x) = \sqrt{1-x},$$

$$B(x) = \sqrt{1-x}\cdot\operatorname{arctanh}(\sqrt{1-x}) = \frac{1}{2}\cdot\sqrt{1-x}\cdot\log\left(\frac{1+\sqrt{1-x}}{1-\sqrt{1-x}}\right).$$

Using the method of variation of parameters, an explicit solution to the ODE is given by $H(x)$ defined as follows:

$$2\sqrt{1-x}\cdot\operatorname{arctanh}(\sqrt{1-x})\log(-1+x) - 2\sqrt{1-x}\left(-\mathrm{Li}_2\left(-\sqrt{1-x}\right) + \mathrm{Li}_2\left(\sqrt{1-x}\right)\right),$$

and, having made suitable choices for the various analytic continuations of these terms, one can write

$$J(x) = H(x) - 2\pi i B(x) + \frac{\pi^2}{2}A(x).$$

In retrospect, an easier (if equivalent) way to write a new $\mathbf{Q}(x)$-linearly independent function (that gives the same span as $J(x)$) while remaining entirely on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ is to consider the integral:

$$\frac{1}{\sqrt{1-x}}\int_0^x \frac{\log(1-t)}{t\sqrt{1-t}}\,dt. \tag{10.1.7}$$

What is surprising in this formulation is the unexpected lack of extra powers of 2 in the denominators of the Taylor series expansion of (10.1.7). While the individual factors $1/\sqrt{1-x}$ and $\int \frac{\log(1-x)}{x\sqrt{1-x}}dx$ have 2-adic convergence discs $|x|_2 < 1/4$ at $x=0$, their product overconverges to the full unit open 2-adic disc $|x|_2 < 1$.

Remark 10.1.8. One can prove that the $k \in \mathbf{N}_{>0}$ for which the Taylor series of

$$\frac{1}{\sqrt{1-x}}\int \frac{\log^{k-1}(1-x)}{x\sqrt{1-x}}dx \tag{10.1.9}$$

converges on the 2-adic unit disc $|x|_2 < 1$ are exactly the positive even integers, and that for these k, the Taylor expansion belongs to

$$J_k(x) := \frac{1}{\sqrt{1-x}}\int \frac{\log^{k-1}(1-x)}{x\sqrt{1-x}}dx \in \sum_{n=1}^{\infty}\frac{x^n}{[1,\ldots,n]^k}\,\mathbf{Z}.$$

This defines a sequence $J_2, J_4, J_6, \ldots$ (with $J_2 = J$) of G-functions holonomic on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$, of denominator types $x^n/[1,\ldots,n]^*$, and independent over the multiple polylogarithm ring § 10.3. △

10.2. **Added integrations.** We define two more functions $B_6(y)$ and $B_7(y)$ as follows:

$$\begin{aligned} B_6(y) &= \int \frac{B_3(y)}{y}\,dy = \int \frac{2\arcsin(\sqrt{y}/2)^2}{y}\,dy = \sum_{n=1}^{\infty} y^n \cdot \frac{(n-1)!^2}{n(2n)!}, \\ B_7(y) &= \int \frac{B_4(y)}{y}\,dy. \end{aligned} \tag{10.2.1}$$

We have:

Lemma 10.2.2. *The denominator types of $B_i(y)$ for $i = 1,\ldots 7$, as defined in equations (10.1.2), (10.1.3), and (10.2.1) are as follows:*

(1) $B_1(y)$ has trivial denominator type.
(2) $B_2(y)$ has denominator type $[1,2,,\ldots,2n]$.
(3) $B_3(y)$ has denominator type $[1,2,\ldots,2n]n$.
(4) $B_4(y)$ has denominator type $[1,2,,\ldots,2n]^2$.
(5) $B_5(y)$ has denominator type $[1,2,,\ldots,2n](2n-1)$, and thus in particular of denominator type $[1,2,\ldots,2n]^2$.
(6) $B_6(y)$ has denominator type $[1,2,\ldots,2n]n^2$, and thus in particular of denominator type $[1,\ldots,n][1,\ldots,2n]n$, and a fortiori *$[1,2,\ldots,2n]^2 n$.*
(7) $B_7(y)$ has denominator type $[1,2,,\ldots,2n]^2 n$.

Proof. This follows in the case of $B_4(y)$ from Lemma 9.0.3, and in the case of $B_7(y)$ from direct integration from the $n = 4$ case. For the remainder, it follows by direct computation since there is an explicit expression in terms of factorials for the general coefficient. □

Remark 10.2.3. In fact the denominators of these functions have a somewhat better type, namely the $[1,\ldots,2n]$ can be relaxed to $n(n-1)\binom{2n}{n}$. In practice this means that the prime product $\prod_{2n/3<p<n} p$ is absent from the $[1,\ldots,2n]$ part of these denominators. This remark seems to not make any improvement in the setup for Theorem A, but the possibility of canceling prime products could be useful to exploit in other contexts. △

We shall prove in § 12 that the seven functions $B_i(y)$ are linearly independent over $\mathbf{Q}(y)$.

10.3. **The multiple polylogarithm ring.** (This section is more of an extended aside and can be omitted on first reading.) Over $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$, a basic construction of G-functions of type $[1,\ldots,n]^\bullet$ is supplied by the single variable multiple polylogarithm functions

$$\mathrm{Li}_{k_1,\ldots,k_d}(x) = \sum_{n_1>n_2>\ldots>n_d} \frac{x^{n_1}}{n_1^{k_1} n_2^{k_2} \cdots n_d^{k_d}}. \tag{10.3.1}$$

Of these, the following eight functions form a maximal $\mathbf{Q}(x)$-linearly independent set with type $n[1,\ldots,n]^2$:

$$1, \quad \mathrm{Li}_1, \quad \mathrm{Li}_{1,1}, \quad \mathrm{Li}_2, \quad \mathrm{Li}_{1,1,1}, \quad \mathrm{Li}_1 \cdot \mathrm{Li}_2, \quad \mathrm{Li}_{1,2}, \quad \mathrm{Li}_3. \tag{10.3.2}$$

In Remark 10.1.8, we found that the multiple polylogarithms do not exhaust all the $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ functions of the type $[1,\ldots,n]^{\bullet}$, and in particular, that we can add to (10.3.2) a ninth independent function (10.1.7) of the $[1,\ldots,n]^2$ type. By symmetrization, these nine $\mathbf{Q}(x)$-linearly independent functions go to nine $\mathbf{Q}(y)$-linearly independent functions on $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$ with $\mathbf{Z}/2$ local monodromies at the elliptic point $y = 4$. However, whereas in Lemma 9.0.3 we proved that the two symmetrization operations $F(x) \rightsquigarrow F^{\pm}(y)$ take the type $[1,\ldots,n]^{\sigma}$ to the type $[1,\ldots,2n]^{\sigma}$, an examination of the polynomials (9.0.7) of the proof reveals that the plus symmetrization F^+ takes the integrated type $n[1,\ldots,n]^{\sigma}$ to the integrated type $n[1,\ldots,2n]^{\sigma}$, but the minus symmetrization F^- spoils the integrated type $n[1,\ldots,n]^{\sigma}$ into $[1,\ldots,2n]^{\sigma+1}$. This is why, as it turns out, the nine $\mathbf{Q}(x)$-independent functions of type $n[1,\ldots,n]^2$ on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ go to only seven $\mathbf{Q}(y)$-independent functions of type $n[1,\ldots,2n]^2$: the above $B_i(y)$. A similar remark shall apply to § 12.1, where under a supposed $\mathbf{Q}$-linear dependency among $1, \zeta(2), L(2,\chi_{-3})$ we would get as many as 17 independent functions over $x \in \mathbf{P}^1 \smallsetminus \{0, 1/9, -1/8, 1, \infty\}$ with the integrated type $n[1,\ldots,n]^2$ and holomorphic at $\{0, 1/9, -1/8\}$ (only the above-listed nine of which really exist), but only 14 of the symmetrizations (seven of them genuine) have the corresponding integrated type $n[1,\ldots,2n]^2$.

One of the key ideas in our paper is that while — as explained in Basic Remark 9.0.20 — our holonomy bounds are equivalent for the data $(\varphi; \prod[1,\ldots,\mathbf{b}n]) := \big(\lambda^2/(\lambda-1), [1,\ldots,2n]^2\big)$ and $(\varphi; \prod[1,\ldots,\mathbf{b}n]) := \big(\lambda, [1,\ldots,n]^2\big)$, the integrated type

$$(\varphi; \prod n^{\mathbf{e}}[1,\ldots,\mathbf{b}n]) := \big(\lambda^2/(\lambda-1), n[1,\ldots,2n]^2\big)$$

yields significantly better bounds than the integrated type

$$(\varphi; \prod n^{\mathbf{e}}[1,\ldots,\mathbf{b}n]) := \big(\lambda, n[1,\ldots,n]^2\big)\,;$$

so much so that the 14 functions in the former type turn out to be a far stronger constraint than the 17 functions in the latter type. We elaborate on this comparison in our next remark.

Remark 10.3.3. (This remark is best appreciated after reading the entire proof of Theorem A, although it still makes the most sense to place it in this section.) The above 17 functions of denominator type $n[1,\ldots,n]^2$ fit into the following refined denominators scheme in Theorem 6.0.2:

$$\mathbf{b} := \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{t}}$$

and the integrations vector

$$\mathbf{e} := (0,1,1,1,1,1,1,1;0;0,0,0,0,1,1,1,1).$$

The first eight entries here are indexed by the row (10.3.2), in precisely this order, where, in view of the term $\max_i(e_i)$ in the definition (6.0.5) of $\tau^{\sharp}$, we opt to subsume Li_2 into the type $n[1,\ldots,n]$ and Li_3 into the type $n[1,\ldots,n]^2$. The ninth entry is the function (10.1.7). Finally, writing $H(x) \in \mathbf{Q}[\![x]\!]$ for the function in

Proposition 11.1.8 below, the eight last entries are the fictive functions

$$H(x),\ H'(x),\ H\left(\frac{x}{x-1}\right),\ H'\left(\frac{x}{x-1}\right),$$
$$\int \frac{H(x)-H(0)}{x}\,dx,\ \int \frac{H\left(\frac{x}{x-1}\right)-H(0)}{x}\,dx,$$
$$\int \frac{H(x)-H(0)}{x-1}\,dx,\ \int \frac{H\left(\frac{x}{x-1}\right)-H(0)}{x-1}\,dx,$$

which turn out to be $\mathbf{Q}(x)$-linearly independent and holonomic on $\mathbf{P}^1 \smallsetminus \{0, 1/9, -1/8, 1, \infty\}$.

Recall that $\tau(\mathbf{b};\mathbf{e}) = \tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e})$ is built out of two pieces. For these denominator types, we calculate

$$\begin{aligned}\tau^\flat(\mathbf{b}) &= \frac{(1+3)\cdot 0 + (5+7)\cdot 1 + (9+11+13+\ldots+33)\cdot 2}{17^2}\\ &= \frac{558}{289} = 1.93079584\ldots\end{aligned}$$

which improves over the crude main denominator cap $\sigma = 2$. The value for $\tau^\flat(\mathbf{b})$ we obtain here is even better than the corresponding value $191/49 = 2\cdot 1.948979\ldots$ that we will use in § 13; for here we can further exploit the special integrated type n of $\mathrm{Li}_1 = \int dx/(x-1)$. But for the other piece $\tau^\sharp(\mathbf{e})$ of $\tau(\mathbf{b};\mathbf{e})$ we get

$$\tau^\sharp(\mathbf{e}) = 83711/242760 = 0.34483\ldots,$$

with the optimal ξ in (6.0.5) being a certain short interval containing the choice $\xi = 57/40$. In total here,

$$\tau(\mathbf{b};\mathbf{e}) = 558/289 + 83711/242760 = 552431/242760 = 2.275626\ldots.$$

This is *very* much inferior to the value $16603/3920 = 2\cdot 2.1173\ldots$ in (13.0.6), and the three additional functions are not nearly enough to compensate, as we now explain.

To look into the numerics of the holonomy quotients, we can choose the map φ as the optimal map of the form

$$\varphi(z) := \lambda(G(z)), \qquad G : (\overline{\mathbf{D}}, 0) \to (\mathbf{D}, 0), \qquad \varphi'(0) = 16G'(0),$$

where concretely G can be (for example) the Riemann mapping for the topological disc inside $\mathbf{D}$ constrained by any simple closed contour that encircles the origin, precisely like the contours we study in § A. To be admissible, the contour must not enclose any of the non-real fiber points in $\lambda^{-1}(1/9)$ and $\lambda^{-1}(-1/8)$, but let us even ignore this point since it will only make the numerics worse. Then the holonomy bound, which would have to compare to $m = 17$, is by the quotient

$$\frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log 16 + \log|G'(0)| - \frac{552431}{242760}}. \tag{10.3.4}$$

Using a (lightly) optimized choice $\mathrm{Gob}(0.92, 110, 23)$ from the gobble contours defined in § A.2, we find that $|G'(0)| = 0.9163768\ldots$ and the quotient (10.3.4) comes out to approximately 22.7527, a rather long distance from the requisite threshold of 17. △

10.3.5. *Perspective on Theorems A and C.* This is why we shall henceforth stick with the type $n[1,\ldots,2n]^2$ functions $f_i(y)$ (such as the above $B_i(y)$), holonomic with singularities at $y=\infty$, at $y=4$ with a $\mathbf{Z}/2$ local monodromy, and with all other singularities being overconvergent for the $f_i(y)$, and close enough to 0. In the application to Theorems A and C, these latter "overconvergent" singularities turn out to be $\{0,-1/72\}$, as we find out in the next section. Armed with Theorem 6.0.2, we will find in § 13 that these singularities $\{0,-1/72\}$ are indeed close enough to 0, and that a holonomy bound smaller than 14 can fortuitously be reached: proving that such 14 independent functions cannot simultaneously exist. Ultimately, this contradicts the supposed $\mathbf{Q}$-linear dependency among $1,\zeta(2)$, and $L(2,\chi_{-3})$, where as many as 7 of the 14 functions arise from any such linear relationship via Lemma 12.1.1.

11. Zagier's sequences A and C

11.1. **Definitions and basic properties.** In this section, we construct a number of holonomic functions converging on the unit disc and extending to holomorphic functions on the universal cover of $\mathbf{P}^1 \setminus \{0,1/9,-1/8,1,\infty\}$, and also on the universal cover of (the orbifold/stack) $\mathbf{P}^1 \setminus \{0,-1/72,4,\infty\}$ where $y=4$ is an elliptic point of order 2. Under the hypothesis that there is a $\mathbf{Q}$-linear relation between 1, $\zeta(2)$, and $L(2,\chi_{-3})$, these functions would have rational coefficients and bounded denominator growth. These constructions all come — in a form very close to what is presented here — from a paper of Zagier [Zag09], but the sequences themselves were certainly considered before then in similar contexts, including in particular in [SB85], and the arguments required to prove the required identities were first observed by Beukers [Beu87]. They arise more or less as solutions to Picard–Fuchs equations associated to modular curves with precisely four cusps. The observation that certain linear combinations of solutions in $\mathbf{Q}[\![x]\!]$ whose coefficients are interesting periods are overconvergent (that is, extend analytically across the singular point of the ODE closest to $x=0$) was exploited by Beukers [Beu87] to give a reinterpretation of Apéry's original proof that $\zeta(2)$ and $\zeta(3)$ are irrational. As Zagier notes, however, the particular functions we consider (associated to the sequences $\mathbf{A}$ and $\mathbf{C}$ in the notation of [Zag09]) — while giving sequences which converge to both $\zeta(2)$ and $L(2,\chi_{-3})$ — "do not converge quickly enough to yield the irrationality of the limit" ([Zag09, p. 360]). To be precise, these simultaneous approximations $u_n/q_n \to L(2,\chi_{-3}), v_n/q_n \to \zeta(2)$ converge at the rate q_n^{-c} where $c = \log 9/\big(2+\log 9\big) = 0.52349\ldots$, whereas, in the classical scheme for irrationality proofs, an exponent $c>1$ would be required. And yet, these functions are precisely the required input in our method to prove the desired irrationality results, by exploiting not simply the convergence properties of these functions on the unit disc $|x|<1$, but also their analytic continuations beyond the boundary point $x=1$.

The following is standard, and can also be read off from [Zag09, Table 3, p. 357]:

Lemma 11.1.1. *The function*

$$x = q\prod_{n=1}^{\infty}\frac{(1-q^n)^4(1-q^{6n})^8}{(1-q^{2n})^8(1-q^{3n})^4} = q - 4q^2 + 10q^3 + \ldots \tag{11.1.2}$$

with $q=e^{2\pi i\tau}$ defines a uniformization map

$$x\,:\,Y_0(6) = \mathbf{H}/\Gamma_0(6) \to \mathbf{P}^1 \setminus \{0,1/9,1,\infty\}$$

taking the $\Gamma_0(6)$ *cusps* $\tau = i\infty, 0, 1/3, 1/2$ *to the respective cusps* $x = 0, 1/9, 1, \infty$.

Note that one can formally invert this power series and write

$$q = x + 4x^2 + 22x^3 + \ldots \in \mathbf{Z}[\![x]\!]. \tag{11.1.3}$$

It follows that any power series in $\mathbf{Z}[\![q]\!]$ can be written formally as a power series in $\mathbf{Z}[\![x]\!]$, and any power series in $\mathbf{Q}[\![q]\!]$ can be written formally as a power series in $\mathbf{Q}[\![x]\!]$.

Let

$$\chi_{-3}(n) = \left(\frac{-3}{n}\right)$$

be the unique primitive character of conductor 3. Consider the theta function of the Eisenstein lattice $\mathbf{Z}[\zeta_3]$:

$$\theta_{-3}(\tau) := \sum_{m,n \in \mathbf{Z}} q^{m^2+mn+n^2}.$$

This is a weight one modular form of level $\Gamma_0(3)$, and incidentally also an Eisenstein series

$$\theta_{-3}(\tau) = 1 + 6\sum_{n=1}^{\infty}\left(\sum_{d|n}\chi_{-3}(n)\right) q^n \in M_1(\Gamma_0(3), \chi_{-3}).$$

On $\Gamma_0(6)$, we get the weight one Eisenstein series

$$\begin{aligned} A := \frac{\theta_{-3}(\tau) + \theta_{-3}(2\tau)}{2} &= 1 + 3\sum_{n=1}^{\infty}\frac{\chi_{-3}(n)q^n}{1-q^n} + 3\sum_{n=1}^{\infty}\frac{\chi_{-3}(n)q^{2n}}{1-q^{2n}} \\ &= 1 + 3q + 3q^2 + 3q^3 + \ldots \in M_1(\Gamma_0(6), \chi_{-3}). \end{aligned}$$

Further we have these weight three Eisenstein series in $M_3(\Gamma_0(6), \chi_{-3})$:

$$\sum_{n=1}^{\infty}\left(\sum_{d|n}(-1)^{d-1}\chi_{-3}(n/d)d^2\right) q^n$$

and

$$\sum_{n=1}^{\infty}\left(\sum_{d|n}\chi_{-3}(d)d^2\right) q^n - \sum_{n=1}^{\infty}\left(\sum_{d|n}\chi_{-3}(d)d^2\right) q^{2n}.$$

Let us write them respectively as $\theta^2 B$ and $\theta^2 C$, where $\theta = (2\pi i)^{-1}d/d\tau = qd/dq$, and the *Eichler integrals* B and C compute to the following:

$$\begin{aligned} B &= \sum_{n=1}^{\infty}\left(\sum_{d|n}(-1)^{d-1}\chi_{-3}(n/d)d^2\right)\frac{q^n}{n^2} = \sum_{n=1}^{\infty}\frac{\chi_{-3}(n)q^n}{n^2(1-q^n)} - 2\sum_{n=1}^{\infty}\frac{\chi_{-3}(n)q^{2n}}{n^2(1-q^{2n})} \\ &= \sum_{n=1}^{\infty}\chi_{-3}(n)n^{-2}\frac{q^n}{1+q^n} = q - \frac{5q^2}{4} + q^3 - \frac{11q^4}{16} + \frac{44q^5}{25} + \ldots, \\ C &= \sum_{n=1}^{\infty}\left(\sum_{d|n}\chi_{-3}(d)d^2\right)\frac{q^n}{n^2} - \frac{1}{4}\sum_{n=1}^{\infty}\left(\sum_{d|n}\chi_{-3}(d)d^2\right)\frac{q^{2n}}{n^2} \\ &= \frac{1}{4}\sum_{n=1}^{\infty}\chi_{-3}(n)\left(4\mathrm{Li}_2(q^n) - \mathrm{Li}_2(q^{2n})\right) = q - q^2 + \frac{q^3}{9} + q^4 - \frac{24q^5}{25} + \ldots \end{aligned}$$

These formulas make it plain[38] that $\lim_{q\to 1} B = \frac{1}{2}L(2,\chi_{-3})$ and $\lim_{q\to 1} C = \frac{1}{4}\zeta(2)$. Moreover, by canceling modularity factors in the opposite weights 1 and -1 (the latter coming from basic properties [Wei77] of Eichler integrals), the forms $A(B - \frac{1}{2}L(2,\chi_{-3}))$ and $A(C - \frac{1}{4}\zeta(2))$ have a weight zero symmetry around the $\Gamma_0(6)$ cusp $\tau = 0$. Expressing these two products in the Hauptmodul coordinate x leads to holonomic functions on $Y_0(6) \cong \mathbf{P}^1 \smallsetminus \{0, 1/9, 1, \infty\}$ (see, for example, [KZ01, § 2.3]) which are overconvergent at $x = 1/9$, and such that the coefficients of the factors AB and AC in $\mathbf{Q}[\![x]\!]$ give rise to simultaneous Apéry limits $\frac{1}{2}L(2,\chi_{-3})$ and $\frac{1}{4}\zeta(2)$ when compared to the coefficients of A (also considered as a function of x). This is Beukers's framework [Beu87] for irrationality proofs. Our next lemma collects these remarks with indications on how to read them off from [Zag09].

Lemma 11.1.4. *Define power series $H_A(x)$, $H_B(x)$, and $H_C(x)$ in terms of the following formulas:*

$$H_A(x) = A(q), \qquad \frac{H_B(x)}{H_A(x)} = B(q), \qquad \frac{H_C(x)}{H_A(x)} = C(q),$$

where $x = x(q)$ is as in Equation 11.1.2, so

$$H_A(x) = 1 + 3x + 15x^2 + 93x^3 + \ldots = \sum a_n x^n,$$
$$H_B(x) = x + \frac{23x^2}{4} + \frac{145x^3}{4} + \frac{3993x^4}{16} + \ldots = \sum b_n x^n,$$
$$H_C(x) = x + 6x^2 + \frac{343x^3}{9} + \frac{788x^4}{3} + \ldots = \sum c_n x^n.$$

Then:

(1) *The functions $H_A(x)$, $H_B(x)$, $H_C(x)$ are multivalued holonomic functions on $\mathbf{P}^1 \smallsetminus \{0, 1/9, 1, \infty\}$; that is, they extend to holomorphic functions on the universal cover.*
(2) *We have $a_n \in \mathbf{Z}$ and $[1,2,\ldots,n]^2 b_n, [1,2,\ldots,n]^2 c_n \in \mathbf{Z}$.*
(3) *The radius of convergence of $H_A(x)$, $H_B(x)$, and $H_C(x)$ is $R = 1/9$. However, any linear combination of the following two functions:*

$$H_B(x) - \frac{L(2,\chi_{-3})}{2} H_A(x), \qquad H_C(x) - \frac{\zeta(2)}{4} H_A(x)$$

has radius of convergence $R = 1$, where

$$L(2,\chi_{-3}) = \sum_{n=1}^{\infty} \frac{\chi_{-3}(n)}{n^2}, \qquad \zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Proof. This result follows from [Zag09, Table 3] and [Zag09, Table 5] (using an argument previously used by Beukers [Beu87]). To orient the reader, note that the sequence a_n is none other than Zagier's sequence $\mathbf{C}$ from [Zag09]. Namely, there is an equality

$$a_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{2k}{k},$$

[38] We have $\mathrm{Li}_2(1) = \zeta(2)$, and the Cesàro regularization $\sum_{n=1}^{\infty} \chi_{-3}(n) := 1/3$ out of the average of the partial sums $1, 0, 0, 1, 0, 0, 1, 0, 0, \ldots$ is the relevant interpretation in this context. This calculation and heuristic are readily made rigorous after an Abel summation.

and a_n satisfies the recurrence

$$(n+1)^2 a_{n+1} - An(n+1)a_n + Bn^2 a_{n-1} = \lambda a_n \tag{11.1.5}$$

for all n ([Zag09, Equation (3)]). Moreover, b_n satisfies the same recurrence (11.1.5) for all $n \neq 0$. In particular, the facts above concerning $H_A(x)$ and $H_B(x)$ are explained in §6 of [Zag09]. On the other hand, c_n satisfies the recurrence

$$(n+1)^2 c_{n+1} - An(n+1)c_n + Bn^2 c_{n-1} = 1 + \lambda c_n \tag{11.1.6}$$

for all $n \geq 0$. While this sequence is not explicitly in [Zag09], it is a disguised form of Zagier's sequence $\mathbf{A}$. More precisely, if one defines the functions

$$\begin{aligned} G_A(x) &= \frac{1}{1+x} \cdot H_A\left(\frac{x}{x+1}\right) = 1 + 2x + 10x^2 + 56x^3 + \ldots \\ G_B(x) &= \frac{1}{1+x} \cdot H_B\left(\frac{x}{x+1}\right) = x + \frac{15x^2}{4} + 22x^3 + \ldots \\ G_C(x) &= \frac{1}{1+x} \cdot H_C\left(\frac{x}{x+1}\right) = x + 4x^2 + \frac{208x^3}{9} + \ldots \end{aligned} \tag{11.1.7}$$

then the coefficients of $G_A(x)$ are exactly Zagier's sequence $\mathbf{A}$, that is, they satisfy equation (11.1.5) except now for the values $(A, B, \lambda) = (7, -8, 2)$, and the coefficients of $G_C(x)$ now satisfy the same recurrence for $n > 0$. This can be easily proved by showing that both functions satisfy the same ODE and then checking that the first few coefficients are in agreement. □

Using this, we deduce the following:

Proposition 11.1.8. *Suppose there exists a* $\mathbf{Q}$*-linear relationship between* 1, $\zeta(2)$, *and* $L(2, \chi_{-3})$, *namely, suppose that*

$$a + b \cdot L(2, \chi_{-3})/2 + c \cdot \zeta(2)/4 = 0$$

for rational numbers a, b, *and* c. *Let*

$$\begin{aligned} H(x) &:= aH_A(x) + bH_B(x) + cH_C(x) \\ &= b\left(H_B(x) - \frac{L(2,\chi_{-3})}{2} H_A(x)\right) + c\left(H_C(x) - \frac{\zeta(2)}{4} H_A(x)\right). \end{aligned} \tag{11.1.9}$$

Then $H(x) \in \mathbf{Q}[\![x]\!]$ *with denominators of shape* $[1, 2, \ldots, n]^2$, *and* $H(x)$ *satisfies the ODE*

$$x(1-x)(1-9x)y'' + (1 - 20x + 27x^2)y' + 3(-1+3x)y = b + \frac{c}{1-x} \tag{11.1.10}$$

Proof. The rationality claims were established above. Either from equation (11.1.5 and (11.1.6) or more directly using the definitions in terms of Eichler integrals following [Beu87], one verifies that $y = H(x)$ satisfies the given differential equation. □

Remark 11.1.11. In [Beu79], Beukers gave alternate proofs of the irrationality of $\zeta(2)$ and $\zeta(3)$ in terms of multiple integrals. For example, Apéry's approximations to $\zeta(2)$ were seen to be coming directly from the integral

$$\iint_{[0,1]^2} \frac{t^n(1-t)^n s^n (1-s)^n}{(1-st)^{n+1}}\, ds dt, \tag{11.1.12}$$

when evaluated as $a_n - b_n\zeta(2)$. We note that the approximations to $L(2,\chi_{-3})$ and $\zeta(2)$ considered here (and in [Zag09]) can also be viewed in the same way. In particular, one can easily verify (either by hand using a little effort or by using [AZ90] without any effort) the identities:

$$L(2,\chi_{-3})H_A(x) - 2H_B(x) = \sum_{n=0}^{\infty} x^n \iint_{[0,1]^2} \frac{9^n s^n t^n (1-s^3)^n (1-t^3)^n}{(1+st+s^2t^2)^{2n+1}} dsdt, \tag{11.1.13}$$

and

$$\zeta(2)G_A(x) - 4G_C(x) = \sum_{n=0}^{\infty} x^n (-1)^n \iint_{[0,1]^2} \frac{(1-s^2)^n(1-t^2)^n}{(1-st)^{n+1}} dsdt. \tag{11.1.14}$$

One easily finds that

$$\max_{[0,1]^2} \left| \frac{9st(1-s^3)(1-t^3)}{(1+st+s^2t^2)^2} \right| = 1, \qquad \max_{[0,1]^2} \left| \frac{(1-s^2)(1-t^2)}{1-st} \right| = 1, \tag{11.1.15}$$

the maxima being obtained at $s = t = 1/2$ in the first case and $s = t = 0$ in the second. It follows that the integrals are all bounded by 1 which gives a transparent proof that the functions $H(x)$ considered in Proposition 11.1.8 overconverge beyond the singularity at $x = 1/9$ to the entire unit disc.

We can also evaluate the geometric series to express the integral formula (11.1.13) as

$$L(2,\chi_{-3})H_A(x) - 2H_B(x) = \iint_{[0,1]^2} \frac{1+st+s^2t^2}{(1+st+s^2t^2)^2 - 9st(1-t^3)(1-s^3)x} dsdt. \tag{11.1.16}$$

By (11.1.15), this formula represents the function $L(2,\chi_{-3})H_A(x) - 2H_B(x)$ as a continuous integral of a family of (rational, as it happens) functions $f_{s,t}(x) \in \mathcal{O}(\mathbf{C} \smallsetminus [1,\infty))$ *holomorphic* on $\mathbf{C} \smallsetminus [1,\infty)$. The property of being a holomorphic function over a complex domain is inherited by any continuous integration over a parameter $(s,t) \in [0,1]^2$, and so the integral representation makes equally transparent the analyticity of (11.1.16) on $\mathbf{C} \smallsetminus [1,\infty)$. This is Zudilin's point of view in [Zud17]. △

Remark 11.1.17. These integral representations, and especially (11.1.16), may also be compared to Zudilin's [Zud03, Riv06, Nes16]

$$\begin{aligned} L(2,\chi_{-4})U(x) - V(x) &:= \iint_{[0,1]^2} \frac{ds\,dt}{\sqrt{(s-s^2)(t-t^2)} \cdot \big(1 - st - (s-s^2)(t-t^2)x\big)} \\ &= -\sum_{n=0}^{\infty} x^n \iint_{[0,1]^2} \frac{(s-s^2)^{n-1/2}(t-t^2)^{n-1/2}}{(1-st)^{n+1}} dsdt \end{aligned}$$

giving rational approximants to Catalan's constant $G = L(2,\chi_{-4})$. In this instance, the integrand peak rate is

$$\max_{[0,1]^2} \left| \frac{(s-s^2)(t-t^2)}{1-st} \right| = \left(\frac{1+\sqrt{5}}{2} \right)^{-5},$$

and indeed the singularities of the linear ODE are at 0, $\left(\frac{1\pm\sqrt{5}}{2}\right)^5$, and ∞: precisely the same as for Apéry's approximants to $\zeta(2)$. Unfortunately, due to the half-integral exponents in this integral representation, the denominators in these rational approximations are as big as $16^n[1,\ldots,2n]^2$.

A different holonomic sequence of rational approximants to G was given by case **E** in [Zag09], where the ODE singularities are $\{0, 1/8; 1/4, \infty\}$ (the first two of which are overconvergent), and the denominator types are $[1,\ldots,n]^2$.

Since $e^2 > 16\cdot(1/4)$ and $e^4 > \left(\frac{1+\sqrt{5}}{2}\right)^5$ (by a wide margin!), this definitely precludes an approach to the irrationality of the Catalan constant by our method using either of these particular families of rational approximants, unless some completely new idea is discovered. $\triangle$

11.2. **The symmetrization of $H(x)$.** Let a, b, and c be complex numbers such that

$$a + b\cdot L(2,\chi_{-3})/2 + c\cdot\zeta(2)/4 = 0.$$

Then we may define $H(x) \in \mathbf{C}[\![x]\!]$ as in equation (11.1.9). If we additionally assume that 1, π^2, and $L(2,\chi_{-3})$ are linearly dependent over $\mathbf{Q}$, then we can choose a, b, and c to be rational, although the arguments of this section will not require this hypothesis.

We now let $G(y)$ be the symmetrization of $H(x)$:

Definition 11.2.1. Let $G(y) = \mathrm{Sym}^+ H(x)$ as defined in equation (9.0.9), so

$$G(y) = H(x) + H\left(\frac{x}{x-1}\right) \in \mathbf{C}[\![y]\!],$$

and let $G_A(x) = \mathrm{Sym}^+ H_A(x)$, so

$$G_A(y) = H_A(x) + H_A\left(\frac{x}{x-1}\right) \in \mathbf{Z}[\![y]\!].$$

Note that $G_A(y) = 2 - 27y + 1014y^2 - 49536y^3 + \ldots$; the function $G_A(y)$ satisfies an order 4 ODE

$$\sum_{i=0}^{3} c_i(y) G_A^{(i)}(y) = 0$$

which we give explicitly later in equation (12.1.5). The span of $G_A(y)$ and its derivatives generates, over $\mathbf{Q}(x)$, the space spanned by $H_A(x)$ and $H_A(x/(x-1))$ and their derivatives (which are both vector spaces of dimension 4). By Lemma 9.0.3 (2), we immediately have the following:

Lemma 11.2.2. *If $a, b, c \in \mathbf{Q}$, then $G(y)$ has denominator type* $[1, 2, \ldots, 2n]^2$.

12. Functional Linear Independence

Let $A_i(x) \in \mathbf{Q}[\![x]\!]$ be a collection of holomorphic functions functions on $\mathbf{P}^1 \setminus S$ for some finite set S. (In our situation, they will all be Siegel G-functions.) Suppose we wish to prove that the $A_i(x)$ are linearly independent over $\mathbf{Q}(x)$ or $\mathbf{C}(x)$. One strategy is as follows. Let γ be a path in $\mathbf{C}$; for example, take a path starting at $x = 0$, avoiding other points in S, and then returning to 0. The functions $A_i(x)$ can be analytically continued along γ, and as we return to $x = 0$ we obtain a sequence of functions $\widehat{A}_i(x)$ which may now have singularities at $x = 0$. Certainly any polynomial relationship between the $A_i(x)$ extends to (the same) polynomial relationship between the $\widehat{A}_i(x)$, and hence also to a polynomial relationship between the $\widehat{A}_i(x)-A_i(x)$, which can sometimes be useful. But we can alternatively consider any identity between the $\widehat{A}_i(x)$ *modulo* functions which are holomorphic at 0. What may (and often does) happen in principle is that this reduces a linear relationship between a large number of functions to a smaller number of functions, and one can hope to employ some form of inductive strategy to establish full linear independence. A typical example is as follows: Suppose that the path γ starts at $0 \in S$ and is a simple loop around a single point $1 \in S$. Then, if a proper subset of the functions $A_i(x)$ are actually holomorphic at $x = 1$, the corresponding $\widehat{A}_i(x)$ vanish modulo holomorphic functions, and we obtain a corresponding linear relationship between the $\widehat{A}_i(x)$ with fewer terms. A basic example of this is as follows. Suppose that

$$A_1(x) = 1, \quad A_2(x) = \log(1-x), \quad A_3(x) = \mathrm{Li}_2(x).$$

After a suitably oriented loop around zero, we have

$$\widehat{A}_1(x) = 1, \quad \widehat{A}_2(x) = \log(1-x) + 2\pi i, \quad \widehat{A}_3(x) = \mathrm{Li}_2(x) + 2\pi i \log(x).$$

Now, modulo holomorphic functions at zero, we obtain a linear relationship between the three functions 0, 0, and $2\pi i \log(x)$. Clearly this forces the coefficient of $\widehat{A}_3(x)$ to be zero, and reduces us to showing that $A_1(x)$ and $A_2(x)$ are linearly independent because $\log(1-x)$ is not a rational function. We will use this strategy a number of times below. Note that another argument in this case would be to consider the functions $\widehat{A}_i(x)-A_i(x)$ which reduces the problem to the $\mathbf{C}(x)$-linear independence of 1 and $\log x$.

Lemma 12.0.1. *The seven functions $B_i(y)$ for $i = 1, \ldots, 7$ defined in Section 10 in equations (10.1.2), (10.1.3), and (10.2.1) respectively are linearly independent over $\mathbf{C}(y)$.*

Proof. We first of all note that the $B_i(y)$ are all elements of $\mathbf{Q}[\![y]\!]$. Therefore any linear dependency over $\mathbf{C}(y)$ upgrades to one over $\mathbf{Q}(y)$, and so it suffices to prove the result over $\mathbf{Q}(y)$.

We begin by proving the linear independence of the $B_i(y)$ for $i = 1, \ldots, 5$. By Lemma 9.0.3, it suffices to prove the linear independence of the 5-functions 1, $\log(1-x)$, $\log^2(1-x)$, $\mathrm{Li}_2(x)$, and and $J(x)$ of equation (10.1.5) over $\mathbf{Q}(x)$. Certainly 1, $\log(1-x)$, and $\log^2(1-x)$ are independent since $\log(1-x)$ is transcendental over $\mathbf{Q}(x)$. These three functions are also defined on $\mathbf{P}^1 \smallsetminus \{1,\infty\}$ which distinguishes then from $\mathrm{Li}_2(x)$ — take a path γ from 0 which winds around $x=1$, then winds around $x=0$, then winds (in the opposite way) around $x=1$, and returns to zero (as in Figure 12.0.7). The first three functions will be invariant, but $\mathrm{Li}_2(x)$ has non-trivial monodromy on this path. So $\mathrm{Li}_2(x)$ is independent of these previous functions. Now suppose that $J(x)$ was a $\mathbf{Q}(x)$-linear combination of $1, \log(1-x), \log^2(1-x), \mathrm{Li}_2(x)$. If the coefficient of $\mathrm{Li}_2(x)$ was non-trivial, then, after scaling, we may assume that it is 1. But now by differentiation, and using the ODE 10.1.6 for $J(x)$, we obtain a new relation over $\mathbf{Q}(x)$ with $1, \log(1-x), \log^2(1-x)$, and $J(x)$ only (with a non-trivial coefficient of $J(x)$ because of the ODE). But $J(x)$ also has non-trivial monodromy over the path γ, so we conclude as for $\mathrm{Li}_2(x)$ above. Now let us return to the $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$ domain and consider the functions $B_6(y)$ and $B_7(y)$. Recall that:

$$B_6(y) = \int \frac{B_3(y)}{y}\, dy, \quad B_7(y) = \int \frac{B_4(y)}{y}\, dy.$$

Using the derivation formula

$$\frac{d}{dy}\left\{A(y)\int F(y)\, dy\right\} = A'(y)\int F(y)\, dy + A(y)F(y), \tag{12.0.2}$$

we can firstly assume that the coefficients of our linear relation in $\mathbf{Q}(y)$ are polynomials, and then by differentiation reduce to an equality of the form

$$a_0 \int \frac{B_3(y)}{y}\, dy + a_1 \int \frac{B_4(y)}{y}\, dy = \sum_{i=1}^{5} b_i(y)B_i(y), \tag{12.0.3}$$

where now a_0 and a_1 are constants which are not both zero and $b_i(y) \in \mathbf{Q}(y)$. We note the

$$\begin{aligned} y(4-y)\frac{d}{dy}B_2(y) &= (2-y)B_2(y) + y^2, \\ y(4-y)\frac{d}{dy}B_3(y) &= -B_2(y) + 2y, \\ y(4-y)\frac{d}{dy}B_4(y) &= (2-y)B_4(y) + (4-y)B_2(y) - 2y(4-y), \\ 2y(4-y)\frac{d}{dy}B_5(y) &= (4-y)B_5(y) - 2B_2(y) + 4y), \end{aligned} \tag{12.0.4}$$

But now let us differentiate (12.0.3) to get

$$a_0\frac{B_3(y)}{y} + a_1\frac{B_4(y)}{y} = \sum_{i=1}^{5} b_i(y)B_i'(y) + b_i'(y)B_i(y). \tag{12.0.5}$$

We see from equation (12.0.4) and equating the coefficients of $B_4(y)$ and $B_3(y)$ respectively (and using the linear independence of $B_i(y)$ for $i = 1, \ldots, 5$ that

$$\begin{aligned} \frac{a_0}{y} &= b_3'(y), \\ \frac{a_1}{y} &= b_4'(y) + \frac{(2-y)}{y(4-y)}b_4(y). \end{aligned} \tag{12.0.6}$$

From the former equation we get (up to constant)

$$b_3(y) = a_0 \log(y),$$

which, since $b_3(y) \in \mathbf{Q}(y)$, can only happen if $a_0 = 0$. We may also write the latter equation as

$$\frac{a_1\sqrt{y(4-y)}}{y} = \frac{d}{dy}\left(b_3(y)\sqrt{y(4-y)}\right)$$

But the integral of the left-hand side is not algebraic (if $a_1 \neq 0$) but the integral of the right-hand side is, so once more this can only happen when $a_1 = 0$, and the linear independence is established. $\square$

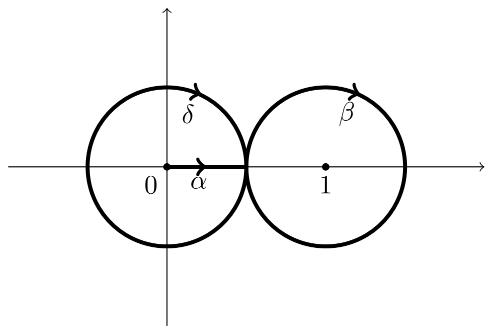

FIGURE 12.0.7. The path $\gamma = \alpha^{-1}\beta^{-1}\delta\beta\alpha$

12.1. **Linear Independence of pure functions and Zagier functions.** Recall from Definition 11.2.1 that

$$G(x) = H(x) + H\left(\frac{x}{x-1}\right).$$

We also let

$$G_A(x) = H_A(x) + H_A\left(\frac{x}{x-1}\right) \in \mathbf{Z}[\![x]\!],$$

which is a homogenous solution to a degree 4 ODE (given explicitly in equation (12.1.5) below). Our final functional linear independence result is as follows:

Lemma 12.1.1 (14 functions)**.** *The seven functions*

$$\int G(y)\,dy, \int \frac{G(y)-G(0)}{y}\,dy, \int \frac{G(y)-G(0)-G'(0)y}{y^2}dy,$$
$$G(y), G'(y), G''(y), G'''(y),$$

together with the seven functions $B_i(y)$ for $i = 1,\ldots,7$, are linearly independent over $\mathbf{C}(y)$.

Remark 12.1.2. It is easy enough to discover Lemma 12.1.1 experimentally. A collection of power series $A_i(x)$ which satisfy a linear relation over $\mathbf{C}(y)$ also satisfy a linear relation with coefficients in $\mathbf{C}[y]$, and thus with coefficients which are polynomials of degree $\leq D$ for some $D \in \mathbf{N}_{>0}$. But the question as to whether there exists such a relation for any given D is equivalent to the vanishing of the determinant of an explicitly computable matrix. Once one establishes that there are no such linear relations for D of moderate size (say $D = 20$), one is sufficiently convinced the result is true and then one writes down a proof. We admit that this is how we arrived at both Lemma 12.1.1 and Lemma 14.3.1, even though there is most likely a higher level proof which better explains the precise numerology. See also Remark 12.1.12. △

Proof. Since the $B_i(y)$ are linearly independent by Lemma 12.0.1, any dependence must include at least one of the terms above with a non-zero coefficient. Let γ denote a path which first traverses 4, then $-1/72$, then 4 in the opposite direction, and then back to 0. The function $G(y)$ is replaced by $\widehat{G}(y)$, which is a solution to the same non-homogenous differential equation at $G(y)$. On the other hand, the functions $\widehat{B}_n(y) = B_n(y)$ remain invariant Since $G(y) \neq \widehat{G}(y)$, we obtain an equivalent relation between the functions

$$\int \widehat{G}(y)\,dy, \int \frac{\widehat{G}(y) - G(0)}{y}\,dy, \int \frac{\widehat{G}(y) - G(0) - G'(0)y}{y^2}\,dy,$$
$$\widehat{G}(y), \widehat{G}'(y), \widehat{G}''(y), \widehat{G}'''(y)$$

and the $B_i(y)$ with the same coefficients. Hence, with $\Delta = \widehat{G}(y) - G(y)$, we obtain a non-zero $\mathbf{C}(y)$-linear relationship between the seven functions

$$\int \Delta(y)\,dy, \int \frac{\Delta(y)}{y}\,dy, \int \frac{\Delta(y)}{y^2}\,dy, \Delta(y), \Delta'(y), \Delta''(y), \Delta'''(y)$$

But $\Delta(y)$ is now a homogenous solution to the corresponding degree 4 ODE which is irreducible, and so by replacing $\Delta(y)$ by its translates under elements of the monodromy group $\pi_1(\mathbf{P}^1 \setminus \{0, -1/72, 4, \infty\})$, we deduce that the corresponding linear relation must hold for any such $\Delta(y)$, including in particular the holomorphic solution $\Delta(y) = G_A(y)$.

Assume such a linear relation exists. After scaling, we may assume that the coefficients lie in $\mathbf{C}(y)$ Using (12.0.2) again:

$$\frac{d}{dy}\left\{A(y)\int F(y)\,dy\right\} = A'(y)\int F(y)\,dy + A(y)F(y), \tag{12.1.3}$$

after repeated differentiation we may assume that the coefficients of the three integral terms are all constants, and that at least one is non-zero. Hence there exists a relation

$$a_0 \int G_A(y)\,dy + a_{-1}\int \frac{G_A(y)}{y}\,dy + a_{-2}\int \frac{G_A(y)}{y^2}\,dy = \sum_{i=0}^{3} b_i(y) G_A^{(i)}(y). \tag{12.1.4}$$

Note that we cannot insist that the $b_i(y) \in \mathbf{C}[y]$, for two reasons. First is that the derivative terms from the integrals involve $G_A(y)$ divided by powers of y. But also when differentiating $G_A^{(3)}(y)$ we obtain $G_A^{(4)}(y)$, and to write this in terms of lower order derivatives in $G_A^{(i)}(y)$ we need to divide by the leading term in the differential equation. In fact, $G_A(x)$ satisfies the ODE

$$\sum_{i=0}^{4} c_i(y) G_A^{(i)}(y) = 0,$$

where $c_i(y)$ are defined as follows:

$$\begin{aligned}
c_0(y) &= -18(3 + 126y - 712y^2 + 360y^3),\\
c_1(y) &= 2(-2 - 2761y + 141632y^2 - 280328y^3 + 176412y^4 - 95616y^5 + 20736y^6),\\
c_2(y) &= 2y(-34 - 6353y + 690355y^2 - 1065613y^3 + 867876y^4 - 438336y^5 + 72576y^6),\\
c_3(y) &= 2(-4 + y)y^2(10 + 204y - 118195y^2 + 146946y^3 - 142848y^4 + 41472y^5),\\
c_4(y) &= (-4 + y)^2 y^3 (1 + 72y)(-1 + 118y - 122y^2 + 144y^3).
\end{aligned} \tag{12.1.5}$$

Thus we can assume that $b_i(y) \in \mathbf{C}[y, c_4(y)^{-1}]$. Differentiating equation (12.1.4) one more time gives an identity

$$a_0 G_A(y) + a_{-1}\frac{G_A(y)}{y} + a_{-2}\frac{G_A(y)}{y^2} = \sum_{i=0}^{3} b_i'(y)G_A^{(i)}(y) + b_i(y)G_A^{(i+1)}(y). \quad (12.1.6)$$

We rewrite (12.1.6) as

$$\begin{aligned} &aG_A(y) + b\frac{G_A(y)}{y} + c\frac{G_A(y)}{y^2} \\ &= \sum_{i=0}^{3} b_i'(y)G_A^{(i)}(y) + \sum_{i=1}^{3} b_{i-1}(y)G_A^{(i)}(y) + b_3(y)G^{(4)}(y), \\ &= \sum_{i=0}^{3} b_i'(y)G_A^{(i)}(y) + \sum_{i=1}^{3} b_{i-1}(y)G_A^{(i)}(y) - \sum_{i=0}^{3} \frac{c_i(y)}{c_4(y)} b_3(y)G^{(i)}(y), \end{aligned} \quad (12.1.7)$$

and thus we deduce the simultaneous equations

$$\begin{aligned} b_3'(y) + b_2(y) - \frac{c_3(y)}{c_4(y)}b_3(y) &= 0, \\ b_2'(y) + b_1(y) - \frac{c_2(y)}{c_4(y)}b_3(y) &= 0, \\ b_1'(y) + b_0(y) - \frac{c_1(y)}{c_4(y)}b_3(y) &= 0, \\ b_0'(y) - \frac{c_0(y)}{c_4(y)}b_3(y) &= a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2}, \end{aligned} \quad (12.1.8)$$

Recall that $b_3(y) \in \mathbf{C}[y, c_4(y)^{-1}]$. Moreover, given $b_3(y)$, one can inductively solve for b_i for $i \in \{2,1,0\}$ from the equation

$$b_i(y) = \frac{c_{i+1}(y)}{c_4(y)}b_3(y) - b_{i+1}'(y).$$

Our strategy is as follows. Since $b_3(y) \in \mathbf{C}(y)$, we can consider the power series expansion of $b_3(y)$ around ∞ and around any point $\alpha \in \mathbf{C}$. Then, by considering the final equation, we obtain an explicit bound on the order of any pole of $b_3(y)$ at α (note that $b_3(y)$ will be holomorphic unless $\alpha = \infty$ or is a root of $c_4(y)$). But that confines $b_3(y)$ to be a rational function such that divisor $(b_3(y)) + D \geq 0$ for an explicit divisor D supported at the roots of $c_4(y)$. This is a finite dimensional (explicitly computable) vector space, and then we can solve for all possible $b_3(y)$ using linear algebra. Another way to view this is to think of this system as a (non-homogenous) ODE in $b_3(y)$, and we are computing the (possible) local expansions around any point using the Frobenius method. Explicitly, with $\alpha \neq \infty$, and

$$b_3(y) = \sum_{i=N}^{\infty} r_i(y-\alpha)^i,$$

(with $N = N_\alpha$ and $r_i - = r_{i,\alpha}$, and suppressing the subscript below) then:

(1) If $\alpha = 0$, the last equality becomes:

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = \frac{-1}{4}(3-N)^2(5-2N)^2 r_N y^{N-4} + \ldots$$

(2) If $\alpha = 4$, the last equality becomes:

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = \frac{-1}{4}(3-N)(2-N)(5-2N)(3-2N)r_N(y-4)^{N-4} + \ldots$$

(3) If $\alpha = -1/72$, the last equality becomes:

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = (3-N)^2(2-N)(1-N)r_N(y+1/72)^{N-4} + \ldots$$

(4) If α is a root β of $144y^3 - 122y^2 + 118y - 1 = 0$, then

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = (3-N)(2-N)(1-N)(1+N)r_N(y-\beta)^{N-4} + \ldots$$

(5) At $\alpha \to \infty$, with

$$b_3(y) = y^N \sum_{i=N}^{\infty} r_i y^{-i},$$

we have

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = \frac{-1}{4}(3-N)^2(5-2N)^2 r_N y^{N-4} + \ldots$$

From this we deduce that:

$$\begin{aligned} N_0 &\geq 2 \\ N_4 &\geq 2 \\ N_{-1/72} &\geq 1 \\ N_\beta &\geq -1 \\ N_\infty &\leq 4. \end{aligned} \tag{12.1.9}$$

From this, it follows that

$$b_3(y) = \frac{y^2(y-4)^2(y+1/72)}{(144y^3 - 122y^2 + 118y - 1)} Q(y), \tag{12.1.10}$$

where $Q(y)$ is a polynomial of degree at most 2. However, if we write

$$Q(y) = q_0 + q_1 y + q_2 y^2,$$

then we find that

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} = \frac{52542464y^{12}q_2 + \ldots}{36y^2(-1 + 118y - 122y^2 + 144y^3)^4} \tag{12.1.11}$$

where the numerator on the right-hand side is a degree 12 polynomial with coefficients linear in $\mathbf{Z}q_0 \oplus \mathbf{Z}q_1 \oplus \mathbf{Z}q_2$. But now by linear algebra one can directly check that there are no choices of the parameters q_i to even make the numerator vanish to order (at least) one at a non-zero root of the denominator. Hence no such $b_3(y)$ exists, and we are done. $\square$

Remark 12.1.12. Suppose instead we had tried to prove the (false!) linear independence of the seven functions $B_i(y)$ together with

$$\int G(y)\,dy, \int \frac{G(y) - G(0)}{y}\,dy, \int \frac{G(y) - G(0) - G'(0)y}{y^2}\,dy,$$
$$\int \frac{G(y) - G(0) - G'(0)y - G''(0)y^2}{y^3}, dy, G(y), G'(y), G''(y), G'''(y),$$

that is, adding another integral. Then the argument would have proceeded exactly as above except now we could only deduce that $N_0 \geq 1$ rather than $N_0 \geq 2$. Then, writing

$$Q(y) = \frac{q_{-1}}{y} + q_0 + q_1 y + q_2 y^2,$$

just as in equation (12.1.11), we would have found that

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} + \frac{a_{-3}}{y^3} = \frac{52542464 y^{13} q_2 + \ldots}{36y^3(-1+118y-122y^2+144y^3)^4}$$

There is now a unique choice of the parameters q_i up to scalar which allows us to remove a single factor of the numerator, namely (up to scalar)

$$Q(y) = \frac{1}{y} - 278 - 844y - 4644y^2.$$

With this choice of $b_3(y)$ as coming from equation (12.1.10), the powers of $(144y^3 - 122y^2 + 118y - 1)$ disappear completely, and we arrive at the equality

$$a_0 + \frac{a_{-1}}{y} + \frac{a_{-2}}{y^2} + \frac{a_{-3}}{y^3} = \frac{676}{9y^2} + \frac{2}{y^3}.$$

This, of course, does now have solutions. This reflects that there *is* a linear dependence between these functions. In fact, there is already a $\mathbf{C}(y)$-linear dependence between the functions

$$\int \frac{G_A(y)}{y^2}\, dy,\ \int \frac{G_A(y)}{y^3}\, dy,\ G_A(y),\ G'_A(y),\ G''_A(y),\ G'''_A(y),\ \text{and } 1.$$

Note as another consistency check with the solution $a_{-2} = 676/9$ and $a_{-3} = 2$, we have (in analogy with equation 12.1.6) that

$$\begin{aligned} a_0 G_A(y) + a_{-1}\frac{G_A(y)}{y} + a_{-2}\frac{G_A(y)}{y^2} + a_{-3}\frac{G_A(y)}{y^3} &= \frac{676}{9}\frac{G_A(y)}{y^2} + 2\frac{G_A(y)}{y^3} \\ &= \frac{4}{y^3} + \frac{866}{9y^2} - \frac{68728}{3} + \ldots \end{aligned}$$

with no $1/y$ term, consistent with it being a derivative of a meromorphic function at $y = 0$. △

13. Proof of the linear independence of $1, \zeta(2)$, and $L(2,\chi_{-3})$

In this section, we complete the proof of Theorem A using the results of Appendix A. The argument is by a contradiction, by proving that a certain G-function cannot exist. Suppose for the contradiction that there exists a $\mathbf{Q}$-linear relation among the periods $1, \zeta(2), L(2,\chi_{-3})$, which we could write as

$$a + b \cdot L(2,\chi_{-3})/2 + c \cdot \zeta(2)/4 = 0 \tag{13.0.1}$$

with some rational integers $a, b, c \in \mathbf{Z}$, not all zero. Proposition 11.1.8 then constructs a certain G-function $H(x) \in \mathbf{Q}[\![x]\!]$ with denominator type $[1,\ldots,n]^2$ and continuing holonomically on $\mathbf{P}^1 \smallsetminus \{0, 1/9, 1, \infty\}$.

Now § 11.2 converts the G-function $H(x)$ to a G-function $G(y) := \mathrm{Sym}^+ H(x) \in \mathbf{Q}[\![y]\!]$ in the symmetrization coordinate

$$y := x + x/(x-1) = x^2/(x-1).$$

Lemma 11.2.2 shows that the denominator type of $G(y) \in \mathbf{Q}[\![y]\!]$ is $[1, \ldots, 2n]^2$. On the other hand, $G(y)$ is holonomic on $y \in \mathbf{P}^1 \smallsetminus \{0, 4, \infty, -1/72\}$, holomorphic on $y \in \mathbf{C} \smallsetminus [4, \infty)$, and with $\mathbf{Z}/2$ local monodromy around $y = 4$.

We apply Theorem 6.0.2 with the 14×2 denominators type array

$$\mathbf{b} := \begin{pmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}^{\mathrm{t}} \tag{13.0.2}$$

and the integrations vector

$$\mathbf{e} := (0, 0, 1; 0, 0, 0, 0, 0, 0; 1, 1, 1, 1, 1),$$

taking over in (6.0.9) after replacing the letter x there by the symmetrization letter

$$y := x + x/(x-1) = x^2/(x-1);$$

and taking the following ordered list of functions $\{f_i\}_{i=1}^{14}$, see (10.1.2), (10.1.4), and (10.2.1):

$$B_1(y),\, B_2(y),\, B_3(y);\, B_4(y),\, B_5(y),\, G(y),\, G'(y),\, G''(y),\, G'''(y),$$
$$B_6(y),\, B_7(y),\, \int G(y)\, dy,\, \int \frac{G(y) - G(0)}{y}\, dy,\, \int \frac{G(y) - G(0) - G'(0)y}{y^2}\, dy.$$

The $\mathbf{Q}(y)$-linear independence of these 14 functions was proved in Lemma 12.1.1. The denominator types were computed in Lemma 10.2.2. For the integrals, we note that the shift in indexing caused by dividing by powers of y means that these functions are not literally of denominator type $n[1, 2, \ldots, 2n]^2$ but rather of type $n[1, 2, \ldots, 2n+3]^2$; this is not an issue by Remark 6.0.12.

Let us denote by $\mathcal{H}_{Y_0(2)}$ the 14-dimensional $\mathbf{Q}(y)$-linear span of these functions. For the analytic maps $\boldsymbol{\varphi}$ figuring in the various holonomy bounds we have developed, we take restrictions $\varphi(rz)$ of the holomorphic mapping $\varphi \in \mathcal{O}(\overline{\mathbf{D}})$ of Lemma A.4.4. From Corollary 9.0.19 used with $\Sigma^0_{Y_0(2)} := \{-1/72\}$, $\Sigma^1_{Y_0(2)} := \emptyset$, $\varphi_{Y_0(2)} := \varphi$, and $U_{Y_0(2)}$ a sufficiently small open neighborhood of the line segment $[-1/72, 0]$, we have the analyticity $\varphi^* \mathcal{H}_{Y_0(2)} \subset \mathcal{M}(\overline{\mathbf{D}})$.

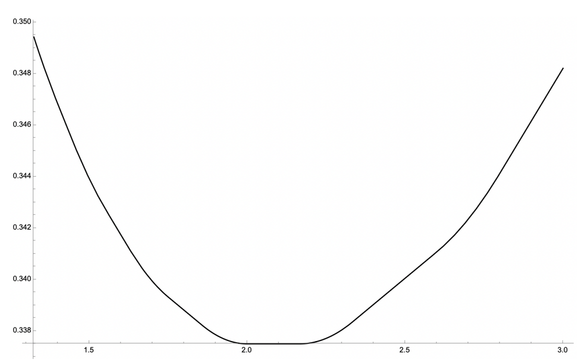

FIGURE 13.0.3. The $\xi \in [1.325, 3]$ fragment of the graph of $(6\xi + I_{\xi}^{14}(\xi))/98$, displaying the interval $\xi \in [2, 13/6]$ as the identical minimizer.

For the denominator rates, we calculate

$$\tau^\flat(\mathbf{b}) = \frac{1\cdot 0 + (3+5)\cdot 2 + (7+9+11+13+\ldots+27)\cdot 4}{14^2} = \frac{191}{49} \qquad (13.0.4)$$

and, from Figure 13.0.3 which reveals $\xi \in [2, 13/6]$ to be the identical minimizer,

$$\begin{aligned}\tau^\sharp(\mathbf{e}) &= \frac{2}{m^2}\min_{\xi\in[0,m]}\left\{\xi\sum_{i=1}^m e_i + \left(\max_{1\le i\le m} e_i\right) I_\xi^m(\xi)\right\} \\ &= \min_{\xi\in[0,14]}\left\{\frac{6\xi + I_\xi^{14}(\xi)}{98}\right\} = \frac{12 + I_2^{14}(2)}{98} = \frac{27}{80}.\end{aligned} \qquad (13.0.5)$$

We obtain

$$\tau(\mathbf{b};\mathbf{e}) = \tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e}) = \frac{191}{49} + \frac{27}{80} = \frac{16603}{3920} = 4.235459\ldots, \qquad (13.0.6)$$

arriving at the number $\frac{191}{49} + \frac{27}{80} = \frac{16603}{3920}$ in (A.5.1).

We can now connect to our holonomy bounds to prove Theorem A. By Proposition 11.1.8 and Lemma 12.1.1, we have a set of $m = 14$ (holonomic) functions linearly independent over $\mathbf{Q}(y)$ that are in $\mathbf{Q}[\![y]\!]$ with the denominator types (13.0.2), contingent upon the $\mathbf{Q}$-linear dependency (13.0.1). Hence it suffices to prove that (any one of) our holonomy bounds yields $m < 14$: this will refute (13.0.1).

For example:

Proof via Theorem 7.0.1. Applying Theorem 7.0.1, we obtain the upper bound $m \le 13.9938\ldots$, as computed in (A.5.1). □

Proof via Theorem 6.0.2. Apply Theorem 6.0.2 with $l = 1, r_0 = e^{-1/2}, \gamma_1 = 14 \cdot 0.209 = 2.926$; we pick this particular parameter based on the numerics in Example 7.6.8. In this case, we have

$$\int_0^1 2t \cdot g^*_{\boldsymbol{\varphi},\boldsymbol{\gamma}}(t)\, dt = 11.316, \ldots$$

and thus the holonomy bound reads

$$m \le \frac{11.316\ldots + \frac{1}{14}\cdot 2.926^2 \cdot \frac{1}{2}}{\log\left(256\cdot\frac{544833945353558660800000000}{865883340756563112243005612 7}\right) - \left(\frac{27}{80} + \frac{191}{49}\right)} = 13.730\ldots < 14.$$

□

Proof via Theorem 7.1.6. With the choice of parameters as in Example 7.4.6 with $r_0 = e^{-1/2}$ and $r_1 = 1$, we obtain the bound (see equation 7.4.7)

$$m \le 13.7206\ldots < 14.$$ □

Of course we may also apply other holonomicity bounds in § 7. See Example 7.4.6 with four parameters r_i rather than two, and Examples 7.5.9 and 7.6.8.

Remark 13.0.7. Had we stayed in the cruder framework $\mathbf{e} = \mathbf{0}$ of Theorem 2.5.1 without added integrals, we would have had to augment $\mathbf{b}$ to the array

$$\mathbf{b}' := \begin{pmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{t}}.$$

Note that although this $\mathbf{b}'$ is not of the particular form in Theorem 6.0.2, we could apply the more general denominator formula (8.0.2) in Theorem 8.0.1. However, for the sake of simplicity, we give a good enough estimate of $\tau^{\flat\flat}(\mathbf{b}')$ which is sufficient to illustrate the necessity of working with nonzero $\mathbf{e}$. On one hand, the upper bound argument in Remark 8.0.6 applies to those $\mathbf{b}'$ such that every column in $\mathbf{b}'$ has two values including one of which is 0. Therefore we have $\tau^{\flat\flat}(\mathbf{b}') \leq (2+2+1) - \frac{1}{14^2}(1^2\cdot 2 + 3^2 \cdot 2 + 8^2 \cdot 1) = \frac{32}{7} = 4.571\ldots$.

On the other hand, by the definition of (8.0.2), we have an easy lower bound by only considering $\mathbf{n}$ satisfying that $n_{j_1} > n_{j_2}$ implies $i_{j_1} \geq i_{j_2}$ and then we have $\tau^{\flat\flat}(\mathbf{b}')$ is at least

$$\frac{1\cdot 0 + 3\cdot 2 + 5\cdot(2+1) + (7+9+\ldots+17)\cdot(2+2) + (19+21+23+25+27)\cdot(2+2+1)}{14^2}$$
$$= \frac{884}{196} = 4.510\ldots,$$

a significantly worse value than (13.0.6). △

14. Products of two logarithms

In this section, we apply our methods to certain products of logarithms. Baker's theorem [Bak22] gives a definitive result for linear forms in logarithms, even over $\overline{\mathbf{Q}}$, but we still do not know how to show that $\log 2 \cdot \log 3$ or $\pi \cdot \log 2$ is irrational. While our methods cannot (as yet!) handle those cases either, we do prove Theorem C, which we recall again here:

Theorem 14.0.1. *Let $m, n \in \mathbf{Z} \smallsetminus \{-1, 0\}$ be integers such that $\left|\frac{m}{n} - 1\right| < \frac{1}{10^6}$. Then*

$$\log\left(1+\frac{1}{m}\right)\log\left(1+\frac{1}{n}\right) \tag{14.0.2}$$

is irrational. Moreover, for $m \neq n$, the following are linearly independent over $\mathbf{Q}$:

$$1, \quad \log\left(1+\frac{1}{m}\right), \quad \log\left(1+\frac{1}{n}\right), \quad \log\left(1+\frac{1}{m}\right)\log\left(1+\frac{1}{n}\right). \tag{14.0.3}$$

Remark 14.0.4. We could certainly improve the constant 10^{-6} by our methods, but some computation suggests that it is unlikely one could do better than (say) 10^{-4}, and most likely not even that far; we make this choice of constant for its relative simplicity. △

The degenerate case of $m = n$ is a trivial consequence of the transcendence of $\log r$ for $r > 0$ in $\mathbf{Q} \smallsetminus \{1\}$, and so we shall assume that $m \neq n$. We begin by recalling a proof of the irrationality of

$$\log\left(1+\frac{1}{m}\right) \tag{14.0.5}$$

for $m \geq 1$ from [AR79, AR80, Chu79, vdP79, vdP80]), based on the method of Apéry limits. It is closely related to the construction we recounted in Basic Remark 2.10.1, and also to the Hermite–Padé construction in § 3.3.7 for the logarithm function.

Let $a > 1$ be an integer. The function

$$A(a,x) := \frac{1}{\sqrt{1-2ax+x^2}} = \sum_{n=0}^{\infty} u_n(a) x^n \tag{14.0.6}$$

lies in $\mathbf{Z}[\![x]\!]$ if a is odd and in $\mathbf{Z}[\![x/2]\!]$ otherwise, and satisfies the first order ODE

$$(1-2ax+x^2)y'+(x-a)y=0. \tag{14.0.7}$$

There is a unique solution to the non-homogenous ODE

$$(1-2ax+x^2)y'+(x-a)y=1$$

with coefficients in $\mathbf{Q}$ that is holomorphic and vanishes at 0; it is given by

$$\begin{aligned} H(a,x) &:= \frac{1}{\sqrt{1-2ax+x^2}}\int_0^x \frac{dt}{\sqrt{1-2at+t^2}} \\ &= \frac{1}{\sqrt{1-2ax+x^2}}\left(\log\left(a-x-\sqrt{1-2ax+x^2}\right)-\log(a-1)\right) \qquad (14.0.8) \\ &= x+\frac{ax^2}{2}+\frac{(3a^2-1)x^3}{6}+\ldots=\sum_{n=0}^{\infty} v_n(a)x^n \in \mathbf{Q}[\![x]\!], \end{aligned}$$

and moreover the coefficients $v_n(a)$ satisfy $[1,2,\ldots,n]v_n(a)\in\mathbf{Z}$ if a is odd and satisfy $[1,2,\ldots,n]2^n v_n(a)\in\mathbf{Z}$ otherwise. By (14.0.8), we have the formula

$$H(a,x)-\frac{1}{2}\log\left(\frac{a+1}{a-1}\right)A(a,x)=\frac{1}{\sqrt{1-2ax+x^2}}\int_{a-\sqrt{a^2-1}}^{x}\frac{dt}{\sqrt{1-2at+t^2}},$$

whose right-hand side overconverges at the singularity $x=a-\sqrt{a^2-1}$ due to multiplying (-1) monodromies of both factors after an analytic continuation along a simple loop enclosing that singularity. This is the same mechanism for overconvergence as in § 2.11.12, as well as in § 11.1 with the canceling automorphy weights in the Eisenstein series A and the Eichler integral $B-\frac{1}{2}L(2,\chi_{-3})$. The case at hand is readily seen to be equivalent, upon notational changes, to the respective ODEs (3.3.12) and formulas (3.3.8) arising from the diagonal Hermite–Padé table for the logarithm function, which we recounted in § 3.3.7 and § 3.3.13.

It follows that

$$\lim_{n\to\infty}\frac{v_n(a)}{u_n(a)}\to\frac{1}{2}\log\left(\frac{a+1}{a-1}\right) \tag{14.0.9}$$

sufficiently quickly to prove the irrationality of this quantity for any odd $a\geq 3$ or any even $a\geq 4$ in light of the inequalities

$$5.828\ldots=3+2\sqrt{2}>e=2.718\ldots,$$
$$7.872\ldots=4+\sqrt{15}>2\cdot e=5.43656\ldots$$

If we let $a=1+2m$, then

$$\log\left(\frac{a+1}{a-1}\right)=\log\left(1+\frac{1}{m}\right),$$

giving the irrationality of (14.0.5), as promised (with a pretty decent irrationality measure, improved further by Chudnovsky [Chu79, Chu83b] by a closer study of this argument).

Now let us consider the arithmetic of the quantities

$$\log\left(\frac{a+1}{a-1}\right)\log\left(\frac{b+1}{b-1}\right) \tag{14.0.10}$$

for pairs of integers $a\neq b$. From (14.0.9), it is obvious that

$$\lim_{n\to\infty}\frac{v_n(a)}{u_n(a)}\cdot\frac{v_n(b)}{u_n(b)}\to\frac{1}{4}\log\left(\frac{a+1}{a-1}\right)\log\left(\frac{b+1}{b-1}\right). \tag{14.0.11}$$

However, this certainly does not converge fast enough to prove irrationality of the right-hand side through any elementary analysis. We shall nevertheless see that as long as a/b is sufficiently close to 1, this quantity is approachable via our new methods based on the function-theoretic properties of the generating series themselves, and basic properties of the Hadamard product operation which allows to construct new G-functions with the desired Apéry limit.

For each pair of integers $a, b \in \mathbf{Z} \setminus \{-1, 0, 1\}$, let us write

$$\eta_a := \frac{1}{2}\log\left(\frac{a+1}{a-1}\right), \quad \eta_b := \frac{1}{2}\log\left(\frac{b+1}{b-1}\right),$$
$$\eta_{a,b} := \eta_a\eta_b := \frac{1}{4}\log\left(\frac{a+1}{a-1}\right)\log\left(\frac{b+1}{b-1}\right).$$

We shall assume that $a \neq \pm b$, as the irrationality and, indeed, the transcendence of $\eta_{a,a} = -\eta_{a,-a} = \eta_a^2$ is already known. Our approach to the arithmetic properties of the product of the Apéry limits $\eta_{a,b} = \eta_a\eta_b$ is via the Hadamard product of the underlying G-functions.

14.1. **Hadamard products and Apéry limits.** Let $\star$ denote the Hadamard product operation on power series: $(\sum a_n x^n) \star (\sum b_n x^n) := \sum a_n b_n x^n$. The function

$$P_A(x) := A(a,x) \star A(b,x) = \sum u_n(a)u_n(b)x^n \tag{14.1.1}$$

satisfies the following ODE $\mathcal{M}(P_A(x)) = 0$:

$$\begin{aligned}
&(-1+x)x(1+x)(1-4abx-2x^2+4a^2x^2+4b^2x^2-4abx^3+x^4)y''\\
&+(-1+8abx+5x^2-12a^2x^2-12b^2x^2+16abx^3-7x^4+4a^2x^4+4b^2x^4-8abx^5+3x^6)y'\\
&+(ab-(-1+3a^2+3b^2)x+8abx^2-(2+a^2+b^2)x^3-abx^4+x^5)y=0.
\end{aligned}\tag{14.1.2}$$

The points $x = 1$ and $x = -1$ are only apparent singularities, as long as $(a - b) \neq 0$ and $(a + b) \neq 0$ respectively. This follows both by general properties [Had1899] of the Hadamard product but can also be verified directly by computing the indicial equation (which is $R(R - 2) = 0$), and then verifying that there are two linearly independent power series solutions. For example, for a putative solution $\sum c_n(x + 1)^n$, the coefficients c_n satisfy a recurrence of the form

$$\begin{aligned}
(a+b)^2 n(n-2)c_n &= 12(a+b)^2 c_{n-1} - 4(a+b)^2(n-2)c_{n-1}(7n-9)(n-2)\\
&\quad + 6(a+b)^2 c_{n-2} + (n-2)c_{n-2}(\ldots) + \ldots,
\end{aligned}$$

which implies that there is at least one solution of the form $x^2 + \ldots$, but there is another of the form:

$$1 + \frac{(x+1)}{2} - \frac{(x+1)^3}{4} + \ldots \in \mathbf{Q}[\![a, b, x+1]\!],$$

and the case of $x = 1$ is similar. The four roots of the quartic are exactly the products of the singularities of the order one ODEs, namely

$$(a \pm \sqrt{a^2-1}) \times (b \pm \sqrt{b^2-1}).$$

Basic Remark 14.1.3. If a and b are large and a/b is very close to one, then the four nonzero finite singularities of (14.1.2) are grouped as follows:

(1) One singularity α very close to 0.
(2) Two singularities very close to 1.

(3) One very large singularity ("close to ∞").

In addition to these, 0 and ∞ themselves are also (essential) singularities. We shall construct (assuming a linear relationship over $\mathbf{Q}$ between the quantities (14.0.3)) a function $H \in \mathbf{Q}[\![x]\!]$ with denominator type $\tau = [1,2,3,\ldots,n]^2$, satisfying a non-homogenous version of (14.1.2), and overconvergent beyond α. When considering functions on $\mathbf{P}^1 \setminus \{0,1,\infty\}$ with $\tau = [1,2,\ldots,n]^2$, we are required by § 2.9.5 to choose an auxiliary function $\varphi : \mathbf{D} \to \mathbf{C} \setminus \{1\}$ with $\varphi(0) = 0$ and $\varphi^{-1}(0) = 0$. If we restrict φ to the disc $D(0, 1-\varepsilon)$ for any $\epsilon \in (0, 1/2]$, then the image of φ will avoid a small open ball containing 1 and an open ball containing ∞, and satisfy $\#\varphi^{-1}(\alpha) = 1$. This is the type of setting where our holonomy bounds can be applied, for we can include not only the (presumably non-existent!) functions $H(x)$ and their derivatives, but also the pure functions on $\mathbf{P}^1 \setminus \{0,1,\infty\}$ the we devised in § 10. As a practical matter, our maps φ are of the form $\lambda \circ \psi$ or, in the equivalent $\mathbf{P}^1 \setminus \{0,4,\infty\}$ setting, $h \circ \psi$ for some map $\psi : \overline{\mathbf{D}} \to \mathbf{D}$, which we take as the Riemann map of a suitably chosen domain in $\mathbf{D}$. But even though the function λ for example avoids 1 on $\mathbf{D}$, to avoid the values within ε of 1 requires taking ψ to have significantly smaller conformal radius unless ε is extremely small. For example, if the image of $\psi(\mathbf{D})$ inside $\mathbf{D}$ included the point 3/4, then $\varphi = \lambda \circ \psi$ on $\mathbf{D}$ would already include the value

$$\lambda(3/4) = 0.9999999999999798332\ldots$$

This numerology is ultimately what forces the hypothesis that $|m/n - 1|$ is very small. $\triangle$

14.1.4. *The overconvergent space.* We now consider solutions to a non-homogenous version of the product ODE (14.1.2). Denoting the corresponding differential operator of (14.1.2) by $\mathcal{M}$, then we also have the following identities:

$$\begin{aligned}
\mathcal{M}(H(a,x) \star A(b,x)) &= -b + 3ax - 3bx^2 + ax^3,\\
\mathcal{M}(A(a,x) \star H(b,x)) &= -a + 3bx - 3ax^2 + bx^3,\\
\mathcal{M}(H(a,x) \star H(b,x)) &= -1 + 2x^2 - x^4.
\end{aligned}$$

We further claim that the functions:

$$\begin{aligned}
P_a &:= (H(a,x) - \eta_a A(a,x)) \star A(b,x) = \sum (v_n(a) - \eta_a u_n(a)) u_n(b) x^n,\\
P_b &:= A(a,x) \star (H(b,x) - \eta_b A(b,x)) = \sum u_n(a)(v_n(b) - \eta_b a_n(b)) x^n,\\
P_{ab} &:= H(a,x) \star H(b,x) - \eta_{a,b} A(a,x) \star A(b,x)\\
&= \sum \left(v_n(a) v_n(b) - \eta_{a,b} u_n(a) u_n(b)\right) x^n\\
&= \sum \left\{\eta_a u_n(a) \left(v_n(b) - \eta_b u_n(b)\right) + \eta_b \left(v_n(a) - \eta_a u_n(a)\right) u_n(b)\right\} x^n\\
&+ \sum \left(v_n(a) - \eta_a u_n(a)\right) \left(v_n(b) - \eta_b u_n(b)\right) x^n
\end{aligned} \tag{14.1.5}$$

are overconvergent beyond the smallest cusp $(a - \sqrt{a^2-1})(b - \sqrt{b^2-1})$. This follows from the bounds

$$|v_n(a) - \eta_a u_n(a)| = O((a - \sqrt{a^2-1})^{n(1-\epsilon)}),$$

$$|v_n(b) - \eta_b u_n(b)| = O((b - \sqrt{b^2-1})^{n(1-\epsilon)}),$$

together with

$$|v_n(a)|, |u_n(a)| = O((a+\sqrt{a^2-1})^{n(1+\epsilon)}),$$
$$|v_n(b)|, |u_n(b)| = O((b+\sqrt{b^2-1})^{n(1+\epsilon)}).$$

Another way to express the power series decomposition of P_{ab} in the last two lines of (14.1.5) is as follows:

$$\begin{aligned} P_{ab} &= \eta_a A(a,x) \star (H(b,x) - \eta_b A(b,x)) + \eta_b (H(a,x) - \eta_a A(a,x)) * A(b,x) \\ &\quad + (H(a,x) - \eta_a A(a,x)) \star (H(b,x) - \eta_b A(b,x)) \end{aligned}$$

The general fact [Had1899] that we exploited here is the overconvergence of the Hadamard product of any set of holonomic power series, at least one among which is an overconvergent branch in the sense of § 2.9. This was essentially combined with the Jacobson identity $1-xy = (1-x)+(1-y)-(1-x)(1-y)$, familiar for example from the proof of the nilpotence of the augmentation ideal of the $\mathbf{F}_p$-group ring of a finite p-group.

14.1.6. *Construction of the unlikely G-function.* Assume now and until the end of the proof of Theorem 14.0.1 that there exist integers r_0, r_a, r_b, and r_{ab} not all zero such that

$$r_a\eta_a + r_b\eta_b + r_{ab}\eta_{a,b} = r_0. \tag{14.1.7}$$

Then the linear combination

$$\begin{aligned} P &:= r_aP_a + r_bP_b + r_{ab}P_{ab} \\ &= r_aH(a,x) \star A(b,x) + r_bA(a,x) \star H(b,x) \\ &\quad + r_{ab}H(a,x) \star H(b,x) + r_0A(a,x) \star A(b,x) \\ &= \sum c_n(a,b)x^n \in \mathbf{Q}[\![x]\!] \end{aligned} \tag{14.1.8}$$

is also from the overconvergent space § 14.1.4, but now it has rational coefficients. This is the G-function, contingent upon our absurd hypothesis of a linear dependency (14.1.7), that will ultimately be rejected by our holonomy bounds. The analytic properties of this unlikely function follow from the overconvergence in § 14.1.4; we now collect the arithmetic properties. If both a and b are odd, then $u_n(a), u_n(b) \in \mathbf{Z}$ and moreover both $[1,2,\ldots,n]v_n(a)$ and $[1,2,\ldots,n]v_n(b)$ lie in $\mathbf{Z}$. Therefore, in the Hadamard products construction,

$$[1,2,\ldots,n]^2c_n(a,b) \in \mathbf{Z}.$$

Moreover, with $a = 1+2m$ and $b = 1+2n$, the non-existence of a linear relationship (14.1.7) is exactly the thesis of Theorem C. (The conditions $a,b \in \mathbf{Z}\smallsetminus\{-1,0,1\}$ become $m,n \in \mathbf{Z}\smallsetminus\{-1,0\}$.) Thus to prove Theorem C it suffices to assume the existence of a relationship (14.1.7) and a function $P(x)$ as in (14.1.8), and establish a contradiction.

Definition 14.1.9. With r_a, r_b, and r_{ab} satisfying (14.1.7), let $P(x)$ be defined as in equation (14.1.8), and let

$$G(y) := P(x) + P\left(\frac{x}{x-1}\right) \in \mathbf{Q}[\![y]\!].$$

With $P_A(x) = A(a,x) \star A(b,x)$ as in (14.1.1), let

$$G_A(y) := P_A(x) + P_A\left(\frac{x}{x-1}\right),$$

hence $G_A(y) \in \mathbf{Z}[\![y]\!]$ if a, b are odd, and $G_A(y) \in \mathbf{Z}[1/2][\![y]\!]$ otherwise. △

As in the proof of Theorem A, we will work with the $Y_0(2)$ picture in the dictionary of § 9, and refute the existence of a G-function $G \in \mathbf{Q}[\![y]\!]$ (contingent upon the existence of a relation (14.1.7)), of the denominators type $[1,\ldots,2n]^2$ and "close to" the $\mathbf{P}^1 \smallsetminus \{0,4,\infty\}$ type that we studied in § 10. We record the following properties of G in Proposition 14.1.11 below, after the following definitions:

Definition 14.1.10. Let $y_{a^\pm,b^\pm}$ denote the $y := \pi(x) = x + \dfrac{x}{x-1}$ images of

$$x = \left(a \pm \sqrt{a^2-1}\right)\left(b \pm \sqrt{b^2-1}\right),$$

where all four pairs of signs are being considered. Let $\mathcal{L}$ denote the pushforward of $\mathcal{M}$ under π, so that $\mathcal{L}(G_A(y)) = 0$. △

Recall that our present discussion is conditional on supposing a $\mathbf{Z}$-linear relation (14.1.7). At this point, we make the additional assumption that the integers $a, b \in \mathbf{Z} \smallsetminus \{\pm 1\}$ are odd.

Proposition 14.1.11. *With $a, b \in \mathbf{Z} \smallsetminus \{\pm 1\}$ odd, the functions $G(y) \in \mathbf{Q}[\![y]\!]$ and $G_A(y) \in \mathbf{Q}[\![y]\!]$ of § 14.1.9 have denominator types $[1,\ldots,2n]^2$ and 1, respectively. Moreover, $\mathcal{L}(G_A(y)) = 0$ and $\mathcal{L}(G(y)) \in \mathbf{Q}[y]$ for some non-zero linear differential operator $\mathcal{L}$ over $\mathbf{Q}(y)$ satisfying:*

(1) $\mathcal{L}$ has no singularities besides $y \in \left\{0, 4, y_{a^\pm,b^\pm}, \infty\right\}$.
(2) $\mathcal{L}$ has $\mathbf{Z}/2$ local monodromy around the singularity $y = 4$.

We shall write down $\mathcal{L}$ explicitly in § 14.2 below; the exact form of the polynomial $\mathcal{L}(G(y)) \in \mathbf{Q}[a,b,y]$ can be computed but will not be important.

14.2. **The differential equation $\mathcal{L}(G_A) = 0$.** Before giving the statement and proof of Lemma 14.3.1 (the analog of of Lemma 12.1.1), we shall examine the ODE $\mathcal{L}(G_A) = 0$ in more detail. Because of the length of this computation, it makes more sense to present it separately rather than interweave it with the proof of Lemma 14.3.1, compared to the corresponding facts concerning the Zagier functions which are proved during the proof of Lemma 12.1.1. However, the reader may well want to look ahead to the statement of Lemma 14.3.1 to see where we are going. One can compute from (14.1.2) the following explicit form of $\mathcal{L}$:

$$\mathcal{L}(G_A) = \sum_{i=0}^{4} c_i(y) G_A^{(i)}(y) = 0,$$

where $c_i(y)$ are certain polynomials with

$$\begin{aligned} c_0(y) &= R_{12}(y) \\ c_1(y) &= R_{16,A}(y) \\ c_2(y) &= yR_{16,B}(y) \\ c_3(y) &= y^2(y-4)R_{15}(y) \\ c_4(y) &= (y-4)^2 y^3 R_4(y) R_{10}(y). \end{aligned} \tag{14.2.1}$$

Here $R_d(y)$ denotes an irreducible polynomial of degree d (with respect to y) in $\mathbf{Q}(a,b,y)$, and the subscripts A and B denote that $R_{16,A}$ and $R_{16,B}$ are distinct. The polynomial

$$(1-x)^4 R_4\left(x + \frac{x}{x-1}\right)$$

has $(a \pm \sqrt{a^2-1})(b \pm \sqrt{b^2-1})$ as 4 of its 8 roots, together with the images of these roots under the involution $w(x) = x/(x-1)$. In particular, the roots of $R_4(y)$ are the singularities $y_{a^\pm,b^\pm}$ of $\mathcal{L}$. The polynomial $R_4(y)$ is given explicitly by

$$\begin{aligned} R_4(y) = {} & 1 + 4y - 8a^2y - 4aby - 8b^2y + 16a^2b^2y + 4y^2 - 12a^2y^2 + 16a^4y^2 \\ & + 20aby^2 - 16a^3by^2 - 12b^2y^2 - 16ab^3y^2 + 16b^4y^2 - 8a^2y^3 + 16aby^3 \\ & - 16a^3by^3 - 8b^2y^3 + 32a^2b^2y^3 - 16ab^3y^3 + 4a^2y^4 - 8aby^4 + 4b^2y^4, \end{aligned}$$

Unlike with $R_4(y)$ or the other accompanying powers of y and $(y-4)$ appearing in $c_4(y)$ of (14.2.1), the roots of $R_{10}(y)$ are not genuine singularities of $\mathcal{L}$. More precisely, this is true if the roots of $R_{10}(y)$ are distinct from those of $R_4(y)y(y-4)$, and this will hold under our assumptions by Lemma 14.2.6 and Lemma 14.2.4

proved below. The polynomial $R_{10}(y)$ is given explicitly by

$$\begin{aligned}
&4a^2b^2+9y-27a^2y-102aby+144a^3by-27b^2y+207a^2b^2y-128a^4b^2y\\
&+144ab^3y-160a^3b^3y-128a^2b^4y+64y^2-93a^2y^2+228a^4y^2-284aby^2+740a^3by^2\\
&-992a^5by^2-93b^2y^2-818a^2b^2y^2-16a^4b^2y^2+740ab^3y^2+1208a^3b^3y^2+228b^4y^2-16a^2b^4y^2\\
&+64a^4b^4y^2-992ab^5y^2+164y^3-132a^2y^3-804a^4y^3+432a^6y^3-100aby^3+1218a^3by^3\\
&+2360a^5by^3-1536a^7by^3-132b^2y^3-1108a^2b^2y^3-856a^4b^2y^3+1856a^6b^2y^3+1218ab^3y^3\\
&-3120a^3b^3y^3-448a^5b^3y^3-804b^4y^3-856a^2b^4y^3-16a^4b^4y^3+2360ab^5y^3-448a^3b^5y^3+432b^6y^3\\
&+1856a^2b^6y^3-1536ab^7y^3+168y^4-18a^2y^4-1476a^4y^4+432a^6y^4-108aby^4-840a^3by^4\\
&+688a^5by^4+384a^7by^4-18b^2y^4+4384a^2b^2y^4-8864a^4b^2y^4+4384a^6b^2y^4-840ab^3y^4\\
&+15744a^3b^3y^4-11744a^5b^3y^4-1476b^4y^4-8864a^2b^4y^4+13920a^4b^4y^4+688ab^5y^4-11744a^3b^5y^4\\
&+432b^6y^4+4384a^2b^6y^4+384ab^7y^4+32y^5+180a^2y^5+2467a^4y^5+792a^6y^5-720a^8y^5\\
&-424aby^5-4228a^3by^5+440a^5by^5+1824a^7by^5+180b^2y^5+3554a^2b^2y^5\\
&-15928a^4b^2y^5+5104a^6b^2y^5-4228ab^3y^5+29392a^3b^3y^5-25088a^5b^3y^5+2467b^4y^5\\
&-15928a^2b^4y^5+37760a^4b^4y^5+440ab^5y^5-25088a^3b^5y^5+792b^6y^5+5104a^2b^6y^5+1824ab^7y^5\\
&-720b^8y^5-32y^6-336a^2y^6+258a^4y^6-3840a^6y^6+480a^8y^6+736aby^6+1064a^3by^6\\
&+6680a^5by^6-4320a^7by^6-336b^2y^6-2676a^2b^2y^6+6976a^4b^2y^6+10688a^6b^2y^6+1064ab^3y^6\\
&-19632a^3b^3y^6-12256a^5b^3y^6+258b^4y^6+6976a^2b^4y^6+10816a^4b^4y^6+6680ab^5y^6-12256a^3b^5y^6\\
&-3840b^6y^6+10688a^2b^6y^6-4320ab^7y^6+480b^8y^6-576a^2y^7-1392a^4y^7+4368a^6y^7+1152aby^7\\
&+5312a^3by^7-12848a^5by^7+576a^7by^7-576b^2y^7-7840a^2b^2y^7+12912a^4b^2y^7-1104a^6b^2y^7\\
&+5312ab^3y^7-8864a^3b^3y^7-768a^5b^3y^7-1392b^4y^7+12912a^2b^4y^7+2592a^4b^4y^7-12848ab^5y^7\\
&-768a^3b^5y^7+4368b^6y^7-1104a^2b^6y^7+576ab^7y^7+192a^2y^8+168a^4y^8-960a^6y^8\\
&-384aby^8-864a^3by^8+1568a^5by^8+192b^2y^8+1392a^2b^2y^8+2368a^4b^2y^8-288a^6b^2y^8\\
&-864ab^3y^8-5952a^3b^3y^8+1152a^5b^3y^8+168b^4y^8+2368a^2b^4y^8-1728a^4b^4y^8+1568ab^5y^8\\
&+1152a^3b^5y^8-960b^6y^8-288a^2b^6y^8+160a^4y^9-640a^3by^9+448a^5by^9+960a^2b^2y^9\\
&-1792a^4b^2y^9-640ab^3y^9+2688a^3b^3y^9+160b^4y^9-1792a^2b^4y^9+448ab^5y^9-32a^4y^{10}\\
&+128a^3by^{10}-192a^2b^2y^{10}+128ab^3y^{10}-32b^4y^{10}.
\end{aligned}$$

We also find that

$$\frac{c_3(y)}{c_4(y)}=\frac{d}{dy}\log\left(\frac{(y-4)^3y^5R_4(y)^3}{R_{10}(y)}\right).$$

We compute that the discriminant of $R_{10}(y)$ has the form (up to an element of $\mathbf{Q}^\times$):

$$\begin{aligned}\Delta_y(R_{10}(y))=&(a-b)^{12}(a+b)^6(1+4a^2-4ab)(1+4b^2-2ab)\\&\times(-3+4a^2-4ab+4b^2)\Phi_{14}(a,b)\Phi_{79}(a,b),\end{aligned}\tag{14.2.2}$$

where $\Phi_d(a,b)\in\mathbf{Q}(a,b)$ is irreducible, satisfies $\Phi_d(a,b)=\Phi_d(b,a)$, and is of degree d when considered as a univariate polynomial in either a or b. We also compute the resultant $\mathrm{Res}_y(R_4(y),R_{10}(y))$ to be, up to a non-zero rational scalar; equal to

$$\begin{aligned}&(a-b)^8(a+b)^6(1+4a^2-4ab)(1+4b^2-2ab)\\&\times(-3+4a^2-4ab+4b^2)^2(9+16a^2-40ab+16b^2)\Phi_{26}(a,b).\end{aligned}\tag{14.2.3}$$

It is easy to verify (reduce modulo 2) that none of the quadratic factors vanish for integer $a, b \in \mathbf{Z}$ and any of the Φ above. One strongly suspects that there are no other integral solutions to $\Phi_d(a,b) = 0$ for the other d except for certain degenerate solutions for some of these polynomials when $a = b$ or $a = -b$. The general Siegel theorem [Zan14, § II.I], see also [BG06, Thm. 7.3.9], certainly guarantees that every irreducible nonrational affine algebraic curve has at most a finite number of integral points; and here, as each of the polynomials Φ_{14}, Φ_{26}, and Φ_{79} turns out to have its highest degree homogeneous piece divisible by $ab(a-b)$ (the degrees of these polynomials are, respectively, $18, 34$, and 92), and hence is not proportional to a power of an irreducible polynomial over $\mathbf{Q}$, Runge's method [BG06, § 9.6.5] (see also [Mas16, § 4] for a gentle and practical introduction) provides in principle an exhaustive algorithm to enumerate all the integer solutions of these equations. For our purposes here, since we are studying the pairs (a, b) with $|a| \asymp |b|$, we shall exploit this hypothesis in the sequel as it spares us the routine but grueling task of carrying out these computations.

Lemma 14.2.4. *Assume that $a, b \in \mathbf{Z} \smallsetminus \{1, 0, -1\}$ with $a \neq \pm b$ satisfy one of the following inequalities:*

$$\left|\frac{a}{b} - 1\right| < \frac{1}{2}, \qquad \left|\frac{a}{b} + 1\right| < \frac{1}{2}.$$

Then:

(1) *$R_4(y)$ is irreducible.*
(2) *$R_{10}(y)$ is co-prime to $R_4(y)$. In particular, the resultant (14.2.3) is non-vanishing.*

Proof. The roots of $R_4\left(x + \dfrac{x}{x-1}\right)$ include $(a - \sqrt{a^2-1})(b - \sqrt{b^2-1})$ as a root. Hence, if we show that $\mathbf{Q}(\sqrt{a^2-1})$ and $\mathbf{Q}(\sqrt{b^2-1})$ are distinct non-trivial real quadratic fields, then $R_4(y)$ is absolutely irreducible since it has at least one root of degree 4. The assumptions on a and b certainly imply that $a^2 - 1$ and $b^2 - 1$ are not squares, so $\mathbf{Q}(\sqrt{a^2-1})$ and $\mathbf{Q}(\sqrt{b^2-1})$ are quadratic fields. If they define the same field, then there exist integers $D, X, Y \in \mathbf{Z}$ with D squarefree such that $(a^2 - 1) = X^2 D$ and $(b^2 - 1) = Y^2 D$, and so (a, X) and (b, Y) are solutions to the Pell equation $u^2 - Dv^2 = 1$. Let us consider the case when $a > b > 0$, the proof applies in the other cases *mutatis mutandis*. After checking the small cases explicitly, we may assume that $b \geq 8$ (note that bounding b also bounds a). We deduce that, for positive algebraic integer unit $\varepsilon > 1$ in $\mathbf{Q}(\sqrt{D})$, there is an equality

$$(a + X\sqrt{D}) = \varepsilon(b + Y\sqrt{D}).$$

The left hand side lies in the interval $[2a-1, 2a]$. The right hand side lies in the interval $\varepsilon[2b-1, 2b]$. Hence

$$1 < \varepsilon < \frac{2a}{2b-1} = \frac{a/b}{1 - \frac{1}{2b}} \leq \frac{3/2}{1 - 1/16} = \frac{16}{10}. \tag{14.2.5}$$

On the other hand, any unit $\varepsilon > 1$ of a real quadratic field satisfies

$$\varepsilon \geq \frac{\sqrt{5}+1}{2} > \frac{16}{10},$$

contradicting equation (14.2.5). The first claim follows. Now if $R_{10}(y)$ has a common factor with $R_4(y)$, it must be divisible by $R_4(y)$. However, we may now synthetically divide one polynomial by the other and the four coefficients of the remaining polynomial of degree ≤ 3 must all be zero. But there are no such solutions in a and b to these four equations — already taking the resultant of any two of them gives an explicit polynomial in a with no integers roots in $\mathbf{Z} \smallsetminus \{-1,0,1\}$. $\square$

We also have the following slightly unpleasant calculus exercise:

Lemma 14.2.6. *Assume that* $a, b \in \mathbf{Z} \smallsetminus \{1,0,-1\}$ *satisfy:*

$$0 < \left|\frac{a}{b} - 1\right| < \frac{1}{10^3}.$$

Then $R_{10}(y)$ *is separable. Moreover,* $\mathrm{Res}_y(R_{10}(y), y(4-y))$ *which equals*

$$48a^2b^2(9+16a^2-40ab+16b^2)(-45+80a^2-128ab+80b^2)\Phi_4(a,b)$$

is non-vanishing.

Proof. Let $\varepsilon = 1/1000$. First let us consider $\Delta_y(R_{10}(y))$ as a polynomial where the coefficient a varies while $b \in \mathbf{Z}$ is fixed. From (14.2.2), the only factors which could possibly vanish for $a \in \mathbf{Z}$ are $\Phi_{14}(a,b)$ and $\Phi_{79}(a,b)$. We examine each of these cases in turn. Consider the case of $\Phi_{14}(a,b)$, and with a and b of the same sign. Let $b = a(1+x)$, so $|x| \le \varepsilon$. Then

$$\begin{aligned}\frac{\Phi_{14}(a,a(1+x))}{\Phi_{14}(a,a)x^2a^4} &= Q_4(x) + a^{-2}Q_2(x) + a^{-2}Q_0(x)(ax)^{-2}\\ &\quad + \frac{\Psi_{-1,12}(a)}{a^3(ax)\Phi_{14}(a,a)} + \sum_{i=0}^{12} x^i \frac{\Psi_{i,12}(a)}{a^4\Phi_{14}(a,a))},\end{aligned} \tag{14.2.7}$$

where $\Psi_{i,12}$ for $i = -1,\dots,12$ are explicit polynomials in a of degree at most 12, and $Q_i(x)$ is an explicit polynomial in x with $Q_i(0) \ne 0$. Moreover, $Q_4(0) = 2$ and is bounded below on the interval $x \in [-1/1000, 1/1000]$ by something only very slightly less than 2. Now we exploit the fact that $b \in \mathbf{Z}$ is an integer to deduce that $ax \in \mathbf{Z}$, and so $|ax| \ge 1$. But assuming $|ax| \ge 1$ and $|x| \le 1/1000$, all the other terms in (14.2.7) are clearly of order $O(a^{-2})$ with explicitly computable constants, and so with the naïve triangle inequality bound, the left-hand side does not vanish as soon as a is large enough. To be completely explicit, we find that, for $|a| \ge 1000$,

$$\begin{aligned}|Q_4(x)| &\ge 1.984,\\ |Q_0(x)|, |Q_2(x)| &\le 10^{-5},\\ \left|\frac{\Psi_{-1,12}(a)}{a^3\Phi_{14}(a,a)}\right| &\le 10^{-10},\\ \left|\frac{\Psi_{i,12}(a)}{a^4\Phi_{14}(a,a)}\right| &\le 10^{-10},\ i = 0,\dots,12\end{aligned}$$

from which the non-vanishing of $\Phi_{14}(a,b)$ comfortably follows from equation (14.2.7). For $|a| < 1000$, note that there are no integers b satisfying the assumed inequalities on a and b. Alternatively, for any integer $|a| \le 1000$, one can check that $\Phi(a,b) = 0$ has no integer roots except for $(a,b) = (1,1)$ and $(-1,-1)$. The argument for $\Phi_{79}(a,b)$ is entirely similar. The analogue of (14.2.7) in this case is

$$\begin{aligned}\frac{\Phi_{79}(a,a(1+x))}{a^{24}x^{12}\Phi_{79}(a,a)} = &\ Q_{24}(x)\\ &+\sum_{i=0}^{5} a^{-2-2i}Q_{22-4i}(x)(ax)^{-2i} + a^{-2-2i}Q_{20-4i}(x)(ax)^{-2i-2}\\ &+\sum_{i=1}^{11}\frac{\Psi_{-i,68}(a)}{a^{24-i}\Phi_{79}(a,a)(ax)^i} + \sum_{i=0}^{67}\frac{\Psi_{i,68}(a)}{a^{24}\Phi_{79}(a,a)}x^i\end{aligned} \tag{14.2.8}$$

Where $\Psi_{i,68}(a)$ has degree at most 68, and $\Phi_{79}(a,a)$ has degree 70. Precisely the same argument as above holds (for $|a| \geq 1000$), again with a (very) comfortable margin, namely,

$$\begin{aligned}|Q_{24}(x)| &\geq 173210,\\ a^{-2}|Q_k(x)| &\leq 30,\ 1 \leq k \leq 12,\\ \left|\frac{a\Psi_{-i,68}(a)}{a^{24-i}\Phi_{79}(a,a)}\right| &\leq 10^{-20},\ i = 1,\ldots,11\\ \left|\frac{\Psi_{i,68}(a)}{a^{24}\Phi_{14}(a,a)}\right| &\leq 10^{-20},\ i = 0,\ldots,67.\end{aligned}$$

□

14.3. **Linear Independence of pure functions and functions arising from G.** Now, in a manner similar to § 12.1 (and with corresponding notation!), we have the following analogue of Lemma 12.1.1. (Remark 12.1.2 concerning Lemma 12.1.1 is equally relevant in this case.)

Lemma 14.3.1 (17 functions, logarithmic version)**.** *Assume that a and b satisfy the assumptions of Lemma 14.2.6. Then the ten functions*

$$\int yG(y)\,dy, \int G(y)\,dy, \int \frac{G(y)-G(0)}{y}\,dy, \int \frac{G(y)-G(0)-G'(0)y}{y^2}\,dy,$$

$$\int \frac{G(y)-G(0)-G'(0)y-G''(0)\frac{y^2}{2}}{y^3}\,dy, \int \frac{G(y)-G(0)-G'(0)y-G''(0)\frac{y^2}{2}-G'''(0)\frac{y^3}{6}}{y^4}\,dy,$$

$$G(y), G'(y), G''(y), G'''(y),$$

together with the seven functions $B_i(y)$ for $i = 1,\ldots,7$, are linearly independent over $\mathbf{C}(y)$.

Proof. We proceed exactly as in the proof of Lemma 12.1.1. Namely, using a monodromy argument we replace $G(y)$ by $\widehat{G}(y)$ and then with $\Delta = \widehat{G}(y) - G(y)$ we reduce to having to show that a certain combination of derivatives and integrals of Δ only are linearly independent. However, Δ will now be a homogenous solution to the ODE $\mathcal{L} = 0$, and so it suffices to consider the case $\Delta(y) = G_A(y)$. As in the proof of Lemma 12.1.1, we are reduced to an equation of the form

$$\sum_{i=-4}^{1} a_i \int G_A(y)y^i\,dy = \sum_{i=0}^{3} b_i(y)G_A^{(i)}(y). \tag{14.3.2}$$

which we analyze by considering the local expansions at the singular points of $\mathcal{L}$ described in Proposition 14.1.11.

Here the roots of $R_4(y)$ are genuine singularities of the ODE, whereas the roots of $R_{10}(y)$ are not. Now, writing

$$b_3(y) = \sum_{i=N}^{\infty} r_i (y-\alpha)^i,$$

with $N = N_\alpha$ and $r_i = r_{i,\alpha}$, just as in the proof of Lemma 12.1.1, we have:

$$\begin{aligned} b_3'(y) + b_2(y) - \frac{c_3(y)}{c_4(y)} b_3(y) &= 0, \\ b_2'(y) + b_1(y) - \frac{c_2(y)}{c_4(y)} b_3(y) &= 0, \\ b_1'(y) + b_0(y) - \frac{c_1(y)}{c_4(y)} b_3(y) &= 0, \\ b_0'(y) - \frac{c_0(y)}{c_4(y)} b_3(y) &= \sum_{i=-4}^{1} a_i y^i, \end{aligned} \tag{14.3.3}$$

and inductively solving for $b_i(y)$ the last equality in equation (14.3.3) around various α is as follows:

(1) If $\alpha = 0$, the last equality becomes:

$$\sum_{i=-4}^{1} a_i y^i = \frac{-1}{4}(3-N)^2(5-2N)^2 r_N y^{N-4} + \dots$$

(2) If $\alpha = 4$, it becomes:

$$\sum_{i=-4}^{1} a_i y^i = \frac{-1}{4}(3-N)(2-N)(5-2N)(3-2N) r_N (y-4)^{N-4} + \dots$$

(3) If α is a root β of $R_4(y)$, the last equality becomes:

$$\sum_{i=-4}^{1} a_i y^i = -(3-N)^2(2-N)(1-N) r_N (zy-\alpha)^{N-4} + \dots$$

(4) If α is a root γ of $R_{10}(y)$,

$$\sum_{i=-4}^{1} a_i y^i = (3-N)(2-N)(1-N)(1+N) r_N (y-\alpha)^{N-4} + \dots$$

(5) At $\alpha \to \infty$, with $b_3(y) = y^N \sum_{i=N}^{\infty} r_i y^{-i}$, we have

$$\sum_{i=-4}^{1} a_i y^i = -(5-N)(4-N)^2(3-N) r_N z^{N-4} r_N y^{N-4} + \dots$$

From these we deduce that:

$$\begin{aligned} N_0 &\geq 0 \\ N_4 &\geq 2 \\ N_\beta &\geq 1 \\ N_\gamma &\geq -1 \\ N_\infty &\leq 5. \end{aligned} \tag{14.3.4}$$

This allows us to write:

$$b_3(y) = \frac{(y-4)^2 R_4(y)}{R_{10}(y)} Q(y), \tag{14.3.5}$$

Hence we may write

$$Q(y) = q_0 + q_1 y + q_2 y^2 + \ldots q_9 y^9.$$

We find that, as a ratio of polynomials in y, we have

$$\sum_{i=-4}^{1} a_i y^i = \frac{S_{42}(y)}{y^4 R_{10}^4(y)} \tag{14.3.6}$$

for a polynomial $S_{42}(y) \in \mathbf{Q}(a,b,y)$ of degree 42. Note for degree reasons, this already implies that $a_1 = a_0 = a_{-1} = 0$), Now solving for $q_0, \ldots, q_9$ in order to account for a single factor of $R_{10}(y)$, we obtain a system of 10 linear equations in 10 unknowns. If we take the corresponding determinant of the matrix, we obtain a (symmetric) polynomial in a and b of the form:

$$(a-b)^{98} a^4 b^4 (a+b)^{12} (1+4a^2-4ab)^2 (1+4b^2-4ab)(-3+4a^2-4ab+4b^2)^2$$
$$(9+16a^2-40ab+16b^2)^2(-45+80a^2-128ab+80b^2)^2$$
$$\Phi_4(a,b)\Phi_6(a,b)\Phi_{14}(a,b)\Phi_{26}(a,b)\Phi_{79}(a,b).$$

But each of these irreducibles factor is also a factor of

$$\Delta_y R_{10}(y) \mathrm{Res}_y(R_{10}(y)R_4(y)) \mathrm{Res}_y(R_{10}(y), y(y-4)),$$

which under our assumptions do not vanish by Lemmas 14.2.4 and 14.2.6 respectively. Hence the determinant is non-zero, which means that the $q_i = 0$, but then all the a_i are zero, and there are no linear relationships, as claimed. □

14.4. **Location of the singularities.** As noted in Proposition 14.1.11, the singularities of $\mathcal{L}$ in the $Y_0(2)$ domain away from $0, 4, \infty$ are located at the points $y_{a^\pm, b^\pm}$ of Definition 14.1.10, given by the $y := x^2/(x-1) = x + x/(x-1)$ images of

$$x = \left(a \pm \sqrt{a^2-1}\right)\left(b \pm \sqrt{b^2-1}\right).$$

Let us assume that $\varepsilon < 10^{-6}$, and that

$$\left|\frac{m}{n} - 1\right| < \varepsilon.$$

If m and n are distinct integers, then $1 \le |m-n| < |n|\varepsilon$, so $|m|, |n| \ge \varepsilon^{-1}$. Let $a = 2m+1$ and $b = 2n+1$. An elementary computation shows that, if $\eta = (a+\sqrt{a^2-1})(b-\sqrt{b^2-1})$, that

$$\left|\eta + \frac{\eta}{\eta - 1}\right| > \frac{1}{\varepsilon}.$$

On the other hand, if $\xi = (a-\sqrt{a^2-1})(b-\sqrt{b^2-1})$, and if $a \ge b$, then $\xi < \varepsilon^2/4$, and

$$\left|\xi + \frac{\xi}{\xi - 1}\right| \le \frac{\varepsilon^4}{16}.$$

Moreover, $\xi^{-1} > 4/\varepsilon^2$, and

$$\left|\xi^{-1} + \frac{\xi^{-1}}{\xi^{-1} - 1}\right| > \frac{16}{\varepsilon^2} > \frac{1}{\varepsilon}.$$

It follows that the singularities of $\mathcal{L}$ are all contained either within the disc

$$D(0,\varepsilon^4/16) = D(0,10^{-12}2^{-4}),$$

or outside the disc

$$D(0,\varepsilon^{-1}) = D(0,10^6).$$

14.5. **The proof of Theorem C.** The overall argument is entirely similar to § 13. A putative $\mathbf{Z}$-linear dependency (14.1.7) with odd integers $a, b \in \mathbf{Z} \smallsetminus \{\pm 1\}$ reduces to the general $\mathbf{Z}$-linear dependency

$$2r_a \log\left(1+\frac{1}{m}\right) + 2r_b \log\left(1+\frac{1}{m}\right) + r_{ab} \log\left(1+\frac{1}{m}\right)\log\left(1+\frac{1}{n}\right) = 4r_0,$$

writing $a = 2m+1, b = 2n+1$. We want to prove that there is no such relation if $0 < |1-m/n| < 10^{-6}$, and we argue for the contradiction. By Proposition 14.1.11, the supposed relation produces a G-function with unlikely analytic and arithmetic properties, including denominator type $[1,\ldots,2n]^2$, which Lemma 14.3.1 promotes to some further associated functions, giving with § 10 a totality of 17 functions of type $n[1,\ldots,2n]^2$ and linearly independent over $\mathbf{Q}(y)$. We are now in a position to reject by this G-function an application of either one among Theorems 6.0.2, 7.0.1, or 7.1.13.

All these theorems are to be used after changing the letter x of their respective statements to the symmetrization letter $y := x^2/(x-1)$, and with the following ordered list $\{f_i\}_{i=1}^{17}$ of 17 functions in Lemma 14.3.1:

$$B_1(y),\ B_2(y),\ B_3(y);\ B_4(y),\ B_5(y),\ G(y),\ G'(y),\ G''(y),\ G'''(y);$$
$$B_6(y),\ B_7(y),\ \int yG(y)\,dy,\ \int G(y)\,dy,\ \int \frac{G(y)-G(0)}{y}\,dy,$$
$$\int \frac{G(y)-G(0)-G'(0)y}{y^2}\,dy,\ \int \frac{G(y)-G(0)-G'(0)y-G''(0)\frac{y^2}{2}}{y^3}\,dy,$$
$$\int \frac{G(y)-G(0)-G'(0)y-G''(0)\frac{y^2}{2}-G'''(0)\frac{y^3}{6}}{y^4}\,dy.$$

See (10.1.2), (10.1.4), and (10.2.1) for the functions $B_1,\ldots,B_7$, and Definition 14.1.9 for the function G (which in the end will not exist). The principal denominator types for this ordered list of functions forms the 17×2 array

$$\mathbf{b} := \begin{pmatrix} 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}^{\mathrm{t}},$$

and the added integrations vector is

$$\mathbf{e} := (0,0,1;0,0,0,0,0,0;1,1,1,1,1,1,1,1).$$

For the ambient analytic map $\varphi \in \mathcal{O}(\overline{\mathbf{D}})$ we now select $\varphi := h\circ\psi$, where — yet again — h is the $Y_0(2)$ hauptmodul written in the $\tau = i\infty$ cusp-filling coordinate $q = e^{2\pi i\tau}$ on the disc by the power series formula (9.0.1), and $\psi : \overline{\mathbf{D}} \to \mathbf{D}$ is the holomorphic mapping from § A.6. Corollary 9.0.19 now applies with Proposition 14.1.11, taking $\Sigma^0_{Y_0(2)} := \{y_{a^-,b^-}\}$ to be the $y := x + w(x) = x^2/(x-1)$

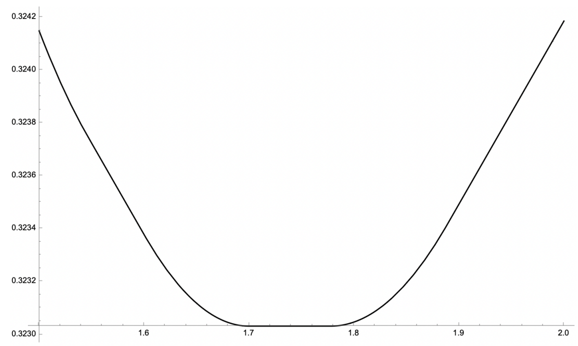

FIGURE 14.5.1. The $\xi \in [1.5, 2]$ fragment of the graph of $(2/17^2)(9\xi + I_\xi^{17}(\xi))$, displaying the interval $\xi \in [1.7, 1.78]$ being contained in the range of minimizer.

image of

$$\begin{aligned}\Sigma^0_{Y(2)} &:= \left\{(a - \sqrt{a^2-1})(b-\sqrt{b^2-1})\right\}\\ &= \left\{\frac{1}{(2n+1+2\sqrt{n^2+n})(2m+1+2\sqrt{m^2+m})}\right\};\end{aligned}$$

and $\Sigma^1_{Y_0(2)} := \emptyset$, $U_{Y_0(2)} := D(0, 1/100)$, and of course, $\varphi_{Y_0(2)} := \varphi = h \circ \psi$. Thus the analyticity conditions for our holonomy bounds are satisfied.

For the denominator rates, the previous calculation now modifies to

$$\tau^\flat(\mathbf{b}) = \frac{1 \cdot 0 + (3+5) \cdot 2 + (7+9+11+13+\ldots+33) \cdot 4}{17^2} = \frac{1136}{289}, \tag{14.5.2}$$

and, from Figure 14.5.1 which reveals $\xi \in [1.7, 1.78]$ to be contained by the minimizing interval,

$$\begin{aligned}\tau^\sharp(\mathbf{e}) &= \frac{2}{m^2} \min_{\xi \in [0,m]} \left\{\xi \sum_{i=1}^m e_i + \left(\max_{1 \le i \le m} e_i\right) I_\xi^m(\xi)\right\}\\ &= \min_{\xi \in [0,17]} \left\{(2/17^2)\left(9\xi + I_\xi^{17}(\xi)\right)\right\}\\ &= (2/17^2)\left(9 \cdot 7/4 + I_{7/4}^{17}(7/4)\right) = \frac{78419}{242760}.\end{aligned} \tag{14.5.3}$$

Hence this time we obtain

$$\begin{aligned}\tau(\mathbf{b};\mathbf{e}) &= \tau^\flat(\mathbf{b}) + \tau^\sharp(\mathbf{e})\\ &= \frac{1136}{289} + \frac{78419}{242760} = \frac{1032659}{242760} = 4.2538\ldots,\end{aligned} \tag{14.5.4}$$

arriving at the number $\frac{1032659}{242760}$ in (A.6.2).

We can once again conclude the proof by deriving a contradiction of the form $m < 17$. Just as in the proof of Theorem A, this can be done in a number of ways. For example: applying Theorem 7.0.1, we obtain the upper bound $m \le 16.2$ as computed in (A.6.2). ☐

15. Complements and further questions

We close our paper with a discussion on some further open problems naturally posed by our method. But first, we discuss more closely the relationship between some of our results and more established methods.

15.1. **Comparison to the Siegel–Bombieri–Chudnovsky theory.**

Theorem C is an irrationality result in two parameters (subject to certain archimedean constraints). In this section, we compare Theorem C to results previously available through the general arithmetic theory of special values of G-functions. We begin by recalling Siegel's definition of a G-function, fixing for this purpose a field embedding $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$:

Definition 15.1.1 (G-function)**.** A power series $f(x) = \sum_{n=0}^{\infty} a_n x^n \in \overline{\mathbf{Q}}[\![x]\!]$ is a *G-function* if it satisfies the following two properties:

(1) $f(x)$ is holonomic: it satisfies an ODE with coefficients in $\mathbf{Q}[x]$.
(2) Both a_n and the denominators of the a_n have moderate growth, namely, the common denominator of $a_0, \ldots, a_n$ grows at most exponentially in n, and the largest Galois conjugate $\boxed{a_n}$ of $a_n \in \overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ grows as most exponentially in n.

From condition (1), it follows that $f(x) \in K[\![x]\!]$ for some number field $K/\mathbf{Q}$. As recalled in § 2.2, one expects ([FR17]) that for such an $f(x)$, there should exist $A \in \mathbf{N}_{>0}$, $b \in \mathbf{Q}_{>0}$, and $\sigma \in \mathbf{N}$ such that

$$a_n A^{n+1} [1, \ldots, bn]^{\sigma} \in \mathcal{O}_K \qquad \forall n \in \mathbf{N}, \tag{15.1.2}$$

and moreover this is known unconditionally under the (conjecturally unnecessary) additional assumption that $f(x)$ arises from geometry [And89, § V app.]. In any case, all the holonomic functions in our paper do have denominators subsumed by (15.1.2); they are manifestly G-functions.

The basic paradigm of the arithmetic theory of G-functions is captured by the following theorem:

Theorem 15.1.3. *For any $\mathbf{Q}(x)$-linearly independent set $f_1, \ldots, f_h \in \mathbf{Q}[\![x]\!]$ of G-functions with* rational *coefficients, there is a constant $N_0 = N_0(f)$, effectively computable from the minimal ODEs of all the f_i, such that the set*

$$\begin{aligned}&\{n \in \mathbf{Z} : \text{the } f_i(1/n) \text{ are } \mathbf{Q}\text{-linearly dependent or contain a divergent value}\}\\ &\subset [-N_0, N_0].\end{aligned} \tag{15.1.4}$$

This result was envisioned in Siegel's 1929 paper [Zan14, § VII] and proved, in the degree of abstraction that we state here, by David and Gregory Chudnovsky [CC85a], after the groundbreaking works of Galočkin [Gal74] (who had to assume the 'factorials canceling property' that reflects in the global nilpotence of the integrable connection; a difficulty already noted by Siegel himself), and Bombieri [Bom81] (who proved a general adelic theorem under the similar and ultimately equivalent condition — but by far easier to check than Galočkin's — that the linear differential system is 'Fuchsian of arithmetic type'.) The Chudnovskys' main result [CC85a, Theorem III], [DGS94, Theorem VIII.1.5], [And89, § VI], [DV01] was precisely the proof of the global nilpotence property for all irreducible ODEs that possess at least one G-series formal solution.

Remark 15.1.5. In line with the discussion in § 3.3.3, the quantitative results on Siegel's program are, of course, stronger and more general than this quintessential form extracted from the works of Galočkin, Bombieri, and the Chudnovskys. See Bombieri's Main Theorem[39] in [Bom81, page 49], and [CC85a, Theorems I and II], [Dèb86, § 1.2 Théorème Principal], [And89, § VII] for other treatments with closely related results. A great picture of the pre-1997 state of the subject is in [PS98, ch. 5 § 7]. For a more recent survey we refer to [Riv19, § 5.6], as well as to [FR18] for further developments. Many of the standard relaxations, such as the admission of the apparently more general special arguments $x = a/n \in \mathbf{Q}$ with $|a/n| < c_1 \exp\left(-c_2\sqrt{\log n \cdot \log\log n}\right)$, can be subsumed into the form (15.1.4) upon making explicit the dependence of N_0 on the differential operator following [Bom81]. But (15.1.4) is also a form that directly connects to integral points on affine algebraic curves, and also to our framework in particular cases § 15.2. We comment on the former connection in our next paragraph.

While Bombieri's general inequality is given in an adelic form over an arbitrary number field, the condition in Theorem 15.1.3 on rational coefficients is of a fundamentally arithmetic nature, and it is crucial for the effectivity clause on N_0. If for instance in the $\{f_1, \ldots, f_h\} := \{1, f\}$ case one wants to handle algebraic number coefficients $f \in \overline{\mathbf{Q}}[\![x]\!]$ like in Remark 8.2.42, the Siegel–Shidlovsky-style proof logic in [CC85a, § 7] based on symmetric powers mandates that the hypothesis $f(x) \notin \overline{\mathbf{Q}}(x)$ (non-rational functions) would need to be strengthened to $f(x) \notin \overline{\mathbf{Q}(x)}$ (transcendental functions). And indeed, Siegel's finiteness theorem on the integral points of non-rational affine algebraic curves has, to this day, not been resolved with an effective upper bound on the heights of the solutions[40], but it can be shown to be equivalent to the $f(x) \in \overline{\mathbf{Q}(x)}$, $\{f_1, \ldots, f_h\} = \{1, f\}$ case of the statement (15.1.4) *with the assumption* $f \in \mathbf{Q}[\![x]\!]$ (of rational coefficients) *relaxed to* $f \in \overline{\mathbf{Q}}[\![x]\!]$ (coefficients from a number field). Hence, a statement such as (15.1.4) is a wide open question for the case of algebraic power series with coefficients from a number field other than $\mathbf{Q}$ or an imaginary quadratic field. The rational coefficients case handled by Theorem 15.1.3 reduces, in the algebraic case of $f(x) \in \overline{\mathbf{Q}(x)} \cap \mathbf{Q}[\![x]\!]$, to Bombieri's extension [BG06, Theorem 9.6.6], [Bom83, § IV], [Dèb85] of the classical *Runge theorem*: an effective resolution in $(x, y) \in \mathbf{Z} \times \mathbf{Q}$ of an irreducible bivariate Diophantine equation $F(x, y) = 0$ over $\mathbf{Q}$ when the highest-order homogeneous part of $F(x, y)$ is not proportional to a power of an irreducible polynomial over $\mathbf{Q}$. (More intrinsically, under the *Runge splitting condition*: the "divisor at infinity" used to give meaning to the integral points problem does not consist of a single Galois orbit of algebraic points on the algebraic curve. As is apparent from the explicit form of Bombieri's inequality [Bom81, page 49], the condition is arithmetic in nature and cannot be attained by extending to a number field; see [BG06, Equation (9.26)] for the general form of Runge's condition over the ring $O_{K,S}$ of S-integers of a number field K.) △

[39] Noting André's remark [And89, page 79] that a scalar coefficient '2' should be added in front of the summation over $\zeta \in \mathrm{sing}_0(L)$ in the term c_{24} in Bombieri's Main Theorem.

[40] This is exactly the content of Hilbert's Tenth problem for the case of Diophantine equations in two variables.

In these optics, our Theorem C may be considered as an $(x, y) = (1/n, 1/m)$ special values analog for the particular *bivariate* G-function

$$F(x,y) := \log(1-x)\log(1-y). \tag{15.1.6}$$

At least some basic results [Gal74, Gal75, Gal96, Hat98, Lys18] of this type are, of course, contained by the single variable Theorem 15.1.3, for instance one can evaluate the univariate G-function $f(x) = \log(1-x)\log(1+x)$ at the point $x = 1/n$. Already on an example as simple as this, the threshold term N_0 arising from the general theory is extremely big; it is estimated in [Hat98] to be on the order of e^{170} in this example. As far as we are aware, the record-lowest threshold on which the irrationality $\log(1-1/n)\log(1+1/n) \notin \mathbf{Q}$ has been proved is Lysov's $n \geq 33$ in [Lys18], by explicit (special!) Hermite–Padé constructions. In this section, we investigate the scope of the general G-function methods on our Theorem C.

15.1.7. *The scope of the single variable theory.* To apply the single variable theory, we should treat $k := m - n$ as a parameter, and consider the G-function

$$f(x) := \log(1-x)\log\left(1 - \frac{x}{1+kx}\right) \in \mathbf{Q}[\![x]\!]; \ k := m - n \in \mathbf{Z}, \tag{15.1.8}$$

whose value at $x = 1/n$ gives the desired product of two logarithms:

$$f(1/n) = \log(1-1/n)\log(1-1/m) = F(1/n, 1/m).$$

For notational simplicity alone, we shall only be concerned here with the irrationality of the product $\log(1-1/n)\log(1-1/m)$, and not with its linear independence from the individual factors; this, of course, suffices for demonstrating the limitations of the general arithmetic theory of special values of G-functions. Thus we apply Theorem 15.1.3 with $\{f_1, \ldots, f_h\} := \{1, f\}$. Then we need to quantify the $N_0 = N_0(k) = N_0(m-n)$ in Theorem 15.1.3 as a function of k, that is essentially of the height of the linear ODE.

We claim that $\log N_0(k) \asymp \log |k|$ for the minimal ODE of the function f_k, by any of methods from the references that we listed in Remark 15.1.5 on the general arithmetic G-function theory.[41] Given this claim, the condition that guarantees $f(1/n) \notin \mathbf{Q}$ from (15.1.4) becomes $|n| > |k|^c$ for some absolute constant $c \in \mathbf{R}_{>0}$, that is the condition

$$|1 - m/n| < |k/n| < |n|^{1-1/c}.$$

This means that the cases of Theorem C that were implicitly known through the general G-functions theory are all under a condition of the form

$$0 < |1 - m/n| \ll |n|^{-\kappa}, \qquad \text{for some } \kappa \in (0,1),$$

a condition necessary and sufficient for these general methods to apply; but a condition significantly stronger than our $0 < |1 - m/n| < 10^{-6}$.

[41] We omit the details, but the enterprising reader can find them in the latex source code for this paper available on the arXiv.

15.1.9. *Multivariable G-function theory.* The arithmetic theory of multivariable G-functions is still in its infancy; we refer to [AB97] for the geometric foundations, and to [Nag97] (for dimension two) and [DV01] (for arbitrary dimension) for the generalizations of the Chudnovskys' fundamental theorem. It seems likely that a two-dimensional version of Siegel's approximating forms scheme § 3.3.1, as carried out by Bombieri in [Bom81], might combine with standard nonvanishing methods [Dys47, Bom82] for Diophantine auxiliary constructions at a special point, to give the irrationality $f(1/n, 1/m) \notin \mathbf{Q}$ of the specializations of a bivariate G-function $f(x, y) \in \mathbf{Q}[\![x, y]\!] \smallsetminus \mathbf{Q}(x, y)$ at arguments of the form $x = 1/n, y = 1/m$, where $n, m \in \mathbf{Z} \smallsetminus \{0\}$ with $\log |n| \gg_f 1$ and $\log |m| / \log |n| \gg_f 1$; to our knowledge, this kind of program has not as yet been worked out in the literature. The range $|1 - m/n| < 10^{-6}$ that we obtained for (15.1.6) from the Apéry limits method is entirely orthogonal to this!

15.1.10. *Avenues from Hermite–Padé constructions.* As explained in §§ 3.3.7, 3.3.13, our proof in § 14 of Theorem C can be conceptually linked to the Hermite–Padé approximants to the logarithm function, used with the Hadamard product construction § 14.1. While products of three or more logarithms appear unreachable by our method here (by the numerology $e^3 > 16$), it could be worthwhile to attempt linear independence of more than a single pair of products of two logarithms, starting from the simultaneous Hermite–Padé approximation theory with several logarithms worked out explicitly in [RT86] and [DHKK22].

A more general scheme, such as we indicated on the most basic examples in § 3.3.3 and § 3.3.7, could be sought with forming the generating function of the special linear forms obtained from evaluating a regular sequence of functional Hermite–Padé approximants to a basic function. Beukers [Beu81, Beu84], using polylogarithms, and Prévost [Pré96], having $\zeta(3, 1 + 1/y)$ for the basic function to be evaluated at the points of the form $y = 1/n$, each were able to interpret the Apéry sequences inside such a scheme. The former type was vastly generalized by Fischler and Rivoal [FR03]. It could be interesting to find a similar interpretation with simultaneous Hermite–Padé approximants for the simultaneous linear forms in $\zeta(2)$ and $L(2, \chi_{-3})$ that we exploited in § 11. We note however that such generating function procedures far from always give rise to G-functions, even if the starting function for the Hermite–Padé approximation is algebraic [BC97b]; for the Prévost type, some non-examples related to zeta values are in [PR21, § 8].

Two other subjects that we have omitted here (in part, for reasons of space) are applications to non-rational algebraic arguments for the two logarithms, as well as p-adic logarithms. Another reason for omitting the later application is that, in this paper, we have emphasized the archimedean place as special when it comes to overconvergence. In [CDT24], we plan to write our holonomicity bound in a more general Arakelov adelic form over a global field.

Finally, speaking more broadly of holonomic explicit constructions by any method, we remark that in many of the more intricate ones in the literature — such as in Zudilin's work [Zud14] on simultaneous approximation to $\zeta(2)$ and $\zeta(3)$, and in Brown and Zudilin's work [BZ22] on $\zeta(5)$ — progress towards a not-yet-attained irrationality goal is measured by setting up a complex set of parameters to maximize the *worthiness exponent* $\limsup\left\{-\frac{\log|\eta-p/q|}{\log q}\right\}$. The latter is very far from faithfully measuring worthiness as a potential for applying in our framework of rational holonomy bounds.

15.2. **Integral holonomic modules.** Although N_0 in Theorem 15.1.3 is effective, it is a wide open[42] problem to precisely (in principle) determine the left-hand side set in (15.1.4). Moreover, as we have seen with a case as simple as [Hat98], the bound on N_0 in the general property (15.1.4) is, in practice, very big. Our findings with the $\mathbf{Q}\left[x,\frac{1}{1-x}\right]$-integrality refinements in §§ 2.7, 2.8, see especially Remarks 2.7.6 and 2.8.2, seem to point towards a completely different approach to those of the cases of (15.1.4) whose holonomicity is recognized by André's arithmetic criterion (Corollary 2.6.1). Our inspiration for hoping to reverse the proof logic in Remark 2.7.6 for the purpose of applying to similar other potential linear independence setups — many of them unproved conjectures — stems from the ideas of Bézivin and Robba [BR89] which they used for reproving the Hermite–Lindemann–Weierstrass theorem as an application of Bertrandias's arithmetic rationality criterion [Ami75, Théorème 5.4.6] (see also [BBR90] for a historical dissection of that proof); and, ultimately, the refinement of those ideas at the hands of André [And00a, And00b] and Beukers [Beu06], using the Chudnovskys's theorem and the Fourier–Laplace duality between E- and G-functions, to reprove (and further refine) the qualitative Siegel–Shidlovsky theorem on special values of E-functions.

Let $\partial := x\cdot(d/dx)$ be the multiplicatively invariant derivation. Consider (for simplicity here) an *étale*[43] holomorphic mapping $\varphi:\mathbf{D}\to\mathbf{C}$ taking $\varphi(0)=0$, and a vector $\mathbf{b}:=(b_1,\ldots,b_r)\in[0,\infty)^r$ with $|\varphi'(0)|>e^{b_1+\cdots+b_r}$. Consider the set $\mathcal{D}$ comprised of the formal power series of the shape

$$f(x)=\sum_{n=0}^{\infty}a_n\frac{x^n}{\prod_{h=1}^{r}[1,\ldots,b_i\cdot n]},\qquad a_n\in\mathbf{Z}\quad\forall n\in\mathbf{N}\tag{15.2.1}$$

such that $\varphi^*f\in\mathcal{O}(\mathbf{D})$ is a holomorphic function on the disc. This is a module over the noncommutative ring $\mathbf{Z}[x,\partial]$: the derivation ∂ acts on the monomials by $\partial(x^n)=nx^n$, which preserves the integrality type in (15.2.1), while the chain rule with the étaleness of φ show that if $f(\varphi(z))$ is holomorphic, so also is

$$\frac{\varphi(z)}{\varphi'(z)}\left(f(\varphi(z))\right)'=\varphi(z)f'(\varphi(z))=(\partial f)(\varphi(z)).$$

By construction, the $\mathbf{Z}[x,\partial]$-module $\mathcal{D}$ is embedded as a submodule of the ring $\mathbf{Q}[\![x]\!]$. Within this ambient ring, $\mathcal{D}$ contains the subring comprised of the $\alpha\in\mathbf{Z}[\![x]\!]$ with $\varphi^*\alpha\in\mathcal{O}(\mathbf{D})$. Let us denote this ring by $\mathcal{O}(\widetilde{\mathcal{V}})$, for reasons related to [BC22] and our Remark 7.3.2, in which this ring is the ring of regular functions on the formal-analytic[44] arithmetic surface we denoted $\widetilde{\mathcal{V}}:=\widetilde{\mathcal{V}}(\varphi)$. Then $\mathcal{D}$ is a module over the ring $\mathcal{O}(\widetilde{\mathcal{V}})$. By Corollary 2.6.1, the field of fractions $\mathrm{Frac}(\mathcal{O}(\widetilde{\mathcal{V}}))$ is a finite

[42]Once again, the exception is the algebraic case $f_i(x)\in\overline{\mathbf{Q}(x)}\cap\mathbf{Q}[\![x]\!]$, in which case a finite computer search is at least in theory enough to finish off this problem in finite computational time.

[43]In other words: the derivative φ' is nowhere vanishing on $\mathbf{D}$.

[44]If φ extends to a holomorphic function on some open neighborhood of the closed disc $\overline{\mathbf{D}}$, to match the convention in [BC22]. For applying the finiteness theorem [BC22, Theorem 9.1.1], this is not a restriction upon considering $\widetilde{\varphi}(z):=\varphi\left((1-\varepsilon)z\right)$ with an $\varepsilon>0$ small enough to still have $|\widetilde{\varphi}'(0)|>e^{b_1+\cdots+b_r}\geq 1$.

field extension of $\mathbf{Q}(x)$, and $\mathcal{D}\otimes_{\mathcal{O}(\widetilde{\mathcal{V}})}\mathrm{Frac}(\mathcal{O}(\widetilde{\mathcal{V}}))$ is a finite-dimensional vector space over that field. Since furthermore $\mathcal{D}$ is preserved by the derivation ∂, there is a finite and $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-stable set of complex algebraic points $\Sigma \subset \overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ such that all elements of $\mathcal{D} \otimes_{\mathcal{O}(\widetilde{\mathcal{V}})} \mathrm{Frac}(\mathcal{O}(\widetilde{\mathcal{V}}))$ (and, *a fortiori*, all elements of $\mathrm{Frac}(\mathcal{O}(\widetilde{\mathcal{V}}))$) continue analytically as *meromorphic* functions along all paths in $\mathbf{C} \smallsetminus \Sigma$.

However, Bost and Charles proved a deeper finiteness theorem [BC22, Theorem 9.1.1]: the ring $\mathcal{O}(\widetilde{\mathcal{V}})$ is a finitely generated $\mathbf{Z}$-algebra. Moreover, their proof leads in principle to an effective algorithm for listing a finite set of generating elements for this $\mathbf{Z}$-algebra. On the other hand, as shown by Remark 2.7.5 on the example $\mathbf{b} = (1)$ and $\varphi(z) = 4z/(1+z)^2$, where $\mathcal{O}(\widetilde{\mathcal{V}}) = \mathbf{Z}[x, 1/(1-x)]$, the $\mathcal{O}(\widetilde{\mathcal{V}})$-module $\mathcal{D}$ is in general infinite. We hence tensor it with $\mathbf{Q}$ and consider $\mathcal{D}_{\mathbf{Q}} := \mathcal{D} \otimes_{\mathbf{Z}} \mathbf{Q}$, which is a torsion-free module over $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}}) := \mathcal{O}(\widetilde{\mathcal{V}}) \otimes_{\mathbf{Z}} \mathbf{Q}$. The latter[45] ring is a finitely generated $\mathbf{Q}$-algebra but also, being of Krull dimension one and integrally closed in its fraction field, it has the added simplicity of being a Dedekind domain. A module over a Dedekind domain is finite and torsion free if and only if it is locally free of finite rank, if and only if it is projective and *generically finite* (where the latter means the finite-dimensionality of the induced vector space over the field of fractions). The content of Theorem 2.7.2 is that, in the previous example of $\mathbf{b} = (1)$ and $\varphi(z) = 4z/(1+z)^2$, the $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})$-module $\mathcal{D}_{\mathbf{Q}}$ is free of rank 2 with basis $\{1, \log(1-x)\}$. However, the content of Remark 2.7.4 is that, in the slightly modified example $\mathbf{b} = (1 + 1/100)$ and $\varphi(z) = 4z/(1+z)^2$ still having $\mathcal{O}(\widetilde{\mathcal{V}}) = \mathbf{Z}[x, 1/(1-x)]$ and $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}}) = \mathbf{Q}[x, 1/(1-x)]$, the $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})$-module $\mathcal{D}_{\mathbf{Q}}$ is infinite while the $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})[1/x]$-module $\mathcal{D}_{\mathbf{Q}}[1/x]$ is once again free of rank 2 with basis $\{1, \log(1-x)\}$. Similar remarks apply to Theorem 2.8.4. In line with these we could ask:

Question 15.2.2. *Can one effectively construct an* $h \in \mathcal{O}(\widetilde{\mathcal{V}}) \smallsetminus \{0\}$ *so that:*

$(\star)$ *The module* $\mathcal{D}_{\mathbf{Q}}[1/h]$ *is locally free over the ring* $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})[1/h]$?

Are there natural verifiable conditions such that this holds even with $h = x$*?*

As remarked above, the rings $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})$ and their localizations are Dedekind domains, and hence, since the $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})[1/h]$-module $\mathcal{D}_{\mathbf{Q}}[1/h]$ is torsion-free and generically finite, the local freeness in $(\star)$ is equivalent to the module being finite, and also to the module being projective. We will see why the insistence on effectivity is the important point for deriving irrationality proofs on $x = 1/n$ special values of certain functions from the holonomic module $\mathcal{D}$. The reason for such a connection is the same as with the André–Beukers (qualitative) refinement [And00a, And00b, Beu06] of the Siegel–Shidlovsky theorem being ultimately derived from a commutative algebra statement formally similar to $(\star)$:

Fact 15.2.3. *The ring of* $\mathbf{E} \subset \mathbf{Q}[\![x]\!]$ *of E-functions with rational coefficients generates over the Laurent polynomial ring* $\mathbf{Q}[x, x^{-1}]$ *an infinite free* $\mathbf{Q}[x, x^{-1}]$*-module* $\mathbf{E}[1/x] = \mathbf{E} \otimes_{\mathbf{Q}[x]} \mathbf{Q}[x, x^{-1}]$.

See [Beu06, Theorem 1.5] for the statement[46] and [Beu06, proof of Cor. 2.2] for the mechanism. The inversion of x here is also necessary, just as we saw with $(\star)$ on

[45] This is a definition in our *ad hoc* notation here, which does not occur in [Bos20] or [BC22].

[46] A theorem of Kaplansky, see [Bos20, § 4.1.2], states that over a Dedekind domain any module which is projective and countably generated, but is not finitely generated, is a free module. Hence

the example $\mathbf{b} = (1+1/100)$ and $\varphi(z) = 4z/(1+z)^2$. In the setting of E-functions, take

$$\{(d/dx)^j \{(e^x - 1)/x\} : j \in \mathbf{N}\}$$

as an example. This set generates an infinite $\mathbf{Q}[x]$-module which localizes to a rank 2 free $\mathbf{Q}[x, x^{-1}]$-module with basis $\{1, e^x\}$. The special role of $x = 0$ in the André–Beukers theory reflects the presence of transcendental E-functions like $f(x) = e^x$ whose minimal ODE does not have 0 for singularity, but yet the special value $f(0) = 1 \in \mathbf{Q}$ is rational.

There is a similar formal mechanism to [BR89, And00b, Beu06] for deriving linear independence proofs if Question 15.2.2 has a positive answer. Suppose $h \in \mathcal{O}(\widetilde{\mathcal{V}}) \smallsetminus \{0\}$ is such that the localized $\mathcal{O}(\widetilde{\mathcal{V}}_{\mathbf{Q}})[1/h]$-module $\mathcal{D}_{\mathbf{Q}}[1/h]$ is locally free. Assume now additionally that, as in § 2.9, there is a contractible open neighborhood $0 \in \Omega \subset \mathbf{D}$ to which φ restricts as a univalent map, and such that $\varphi^{-1}(\Sigma) \subset \Omega$. Consider $f(x) \in \mathcal{D}_{\mathbf{Q}}$ and a finite, $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$-stable set of complex algebraic points $S_f \subset \overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ containing Σ, such that $f(x)$ continues *holomorphically* along all paths in $\mathbf{C} \smallsetminus S_f$. Assume furthermore that those various analytic continuations end up taking on at least two distinct values at the point $x = 1/n$.

Now suppose $n \in \mathbf{Z} \smallsetminus \{0\}$ obeys the following restrictions:

(1) $\varphi^{-1}(1/n) \subset \Omega$;
(2) $1/n \notin S_f$;
(3) All analytic continuations of the element h in $\mathbf{C}^1 \smallsetminus \Sigma$ take nonzero values at the point $1/n$.

Then the special[47] *value $f(1/n)$ satisfies an irrationality property as in (15.1.4): either $f(1/n) \notin \mathbf{Q}$, or else $f(1/n)$ is divergent.*

The point is formally the same as in [Beu06]. There is a non-zero polynomial $Q_f \in \mathbf{Z}[x] \smallsetminus \{0\}$ such that $\{Q_f = 0\} = S_f$ and $Q_f(x)f(x)$ continues *holomorphically* along all paths in $\mathbf{C} \smallsetminus \Sigma$. If $f(1/n) = p/q$ were rational, the local univalence property (1) in our setup from § 2.9 would apply to the function

$$\widetilde{f}(x) := Q_f(x)\frac{f(x) - f(1/n)}{1 - nx} \in \mathbf{Q}[\![x]\!],$$

with $\Sigma^1 := \{s \in \Sigma : \varphi^{-1}(s) = \emptyset\}$ and $\Sigma^0 := \{1/n\} \cup (\Sigma \smallsetminus \Sigma^1)$ and Σ of Proposition 2.9.3 augmented by $\Sigma \cup \{1/n\}$, to derive that

$$\partial^j \left\{\widetilde{f}(x)\right\} \in \mathcal{D}_{\mathbf{Q}}, \qquad \text{for all } j \in \mathbf{N}. \tag{15.2.4}$$

The assumptions ($\star$), (2), and (3) imply that the functions (15.2.4) generate a finite $\mathbf{C}(x)$-module, as well as a finite $\mathbf{C}[\![\frac{1}{1-nx}]\!]$-module. Hence $\mathcal{L}(\widetilde{f}) = 0$ for some nonzero linear differential operator $\mathcal{L}$ over $\mathbf{C}(x)$ which is non-singular at the point $x = 1/n$. But this conflicts with our condition that $\widetilde{f}$ has some analytic continuation $\widetilde{F}$, necessarily also a solution of the ODE $\mathcal{L}(\widetilde{F}) = 0$, such that $x = 1/n$ is a meromorphic pole of $\widetilde{F}(x)$.

---

the $\mathbf{Q}[x, x^{-1}]$-freeness of $\mathbf{E}[1/x]$ reduces to the freeness of all its finitely generated $\mathbf{Q}[x, x^{-1}]$-submodules, that is to [Beu06, Theorem 1.5].

[47] This is meant as $\mathbf{Q}[\![x]\!]$ series evaluated at $x = 1/n$; hence the usual dichotomy with "either irrational or divergent." We can, however, say here the more precise conclusion of irrationality of the value $f(1/n) \in \mathbf{C}$, well-defined by analytic continuation from $x = 0$ staying within the univalent leaf $\Omega \supset \{0, 1/n\}$. An inclusion into a Ball–Rivoal framework of special values beyond the disc of convergence for certain G-functions has been recently achieved by Fischler and Rivoal [FR21].

This puts some prize on constructing an element $h \in \mathcal{O}(\widetilde{\mathcal{V}}) \smallsetminus \{0\}$ for the finiteness property ($\star$) in Question 15.2.2. As a simple example, if $h = 1$ or even just $h = x$ (analogously to the André–Beukers Fact 15.2.3 and the discussion preceding it) is admissible for the bivalent map $\varphi(z) = 8(z+z^3)/(1+z)^4$ of Basic Remark 2.11.1 and the $[1,\dots,n]^2$ denominator type ($\mathbf{b} = (1,1)$), that by itself would suffice — in lieu of the full Conjecture 2.8.1 in that context, which could be more difficult — to embed Remark 2.8.2 into the above discussion, and conclude at one stroke the irrationalities $\mathrm{Li}_2(1/n) \notin \mathbf{Q}$ for the remaining values $n \in \{-4,-3,-2,2,3,4,5\}$. Indeed, on this example, an easy computation (see, for example, [Rob68]) gives $\mathcal{O}(\widetilde{\mathcal{V}}) = \mathbf{Z}[x, 1/(1-x)]$ with $\Sigma = \{0,1,\infty\}$, and for Ω we can take any open neighborhood of $(-1,1)$ in $\mathbf{D}$ that is small enough to have $\varphi^{-1}(\varphi(\Omega)) = \Omega$.

15.3. **Quantitative aspects of linear independence.** As is usual[48] in transcendental number theory, our linear independence proofs in this paper can in principle by promoted to quantitative lower bounds on the linear forms in the relevant periods. In this case, however, the transition is not straightforward and requires a substantial amount of added work that we decided to not engage with in the present paper. The discussion in § 3.3.3 points to a first methodological clue for making such a transition. We plan to turn to this in a future work.

15.4. **The structure ring.** It would be interesting to clarify the scope of arithmetic characterization theorems of the kind of Theorem 2.7.2 (on $\log x$) and Theorem 2.8.4 (on $\log^2 x$), and of the more precise Conjecture 2.8.1 on the $[1,\dots,n]^2$ layer G-functions on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$. One could for example ask how much of the multiple polylogarithm ring § 10.3 may be captured in arithmetic algebraization terms. The following question falls short of our methodology:

Question 15.4.1. *Consider $\mathcal{H}$ to be the $\mathbf{Q}(x)$-vector space generated by functions $f(x)$ of the form*

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\dots,n]^3} \in \mathbf{Q}[\![x]\!], \qquad a_n \in \mathbf{Z} \quad \forall n \in \mathbf{N} \tag{15.4.2}$$

arising from G-functions of geometric origin on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$. Is $\mathcal{H}$ finite dimensional?

The universal map $\varphi : \mathbf{D} \to \mathbf{C} \smallsetminus \{1\}$ taking $\varphi(0) = 0$ and satisfying $\varphi^{-1}(0) = \{0\}$ is λ. Since $16 < e^{\tau} = e^3$, our methods have nothing directly to say about Question 15.4.1; we do not even have a guess as to what the answer might be. (To contrast, for the denominator types — for example — $[1,\dots,n]^2[1,\dots,n/2]$ or $\prod_{k=1}^{8}[1,\dots,n/k]$, Corollary 2.6.1 proves that the corresponding G-functions on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ form a finite-dimensional space, although it is probably quite difficult to determine these spaces, even conjecturally.)

We emphasize that the problem does not necessarily become any easier even when $\tau = 0$; one can ask for which $\alpha \in \mathbf{Q}$ the $\mathbf{Q}(x)$-vector space generated by *algebraic* power series $f(x)$ in $\mathbf{Z}[\![x]\!]$ on $\mathbf{P}^1 \smallsetminus \{0,\alpha,\infty\}$ is infinite. This space is finite when $\alpha > 1/16$, and infinite when $\alpha = 1/16$, where one can construct such functions by writing modular functions with integer coefficients (on congruence subgroups) in terms of $x = \lambda/16$. The main result of [CDT21] was to show that all such $f(x)$

[48] *Except* for André's *transcendance sans transcendance* [And00a, And00b] applying the arithmetic theory of G-functions to recover the Siegel–Shidlovsky theorem on E-functions.

arise in this way. But as soon as $\alpha < 1/16$, we have no methods to understand this problem, or even to determine whether there exists a single non-rational function in $\mathcal{H}$. One can, however, leverage the $\alpha = 1/16$ example to show that for certain templates with $\tau > 0$, the space of G-functions is infinite dimensional.

Proposition 15.4.3. *Let $\mathcal{H}$ denote the $\mathbf{Q}(x)$-vector space generated by functions $f(x)$ of the form*

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1,\ldots,2n]^2} \in \mathbf{Q}[\![x]\!], \qquad a_n \in \mathbf{Z} \quad \forall n \in \mathbf{N} \tag{15.4.4}$$

arising from G-functions of geometric origin on $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$. Then $\mathcal{H}$ is infinite dimensional.

Proof. Let $g(q) \in \mathbf{Z}[\![q]\!]$ be a modular function on $X_0(N)$ which is holomorphic away from the cusps. Then, writing $f(q)$ as a function of $x = \lambda/16$, we find that

$$g(x) \in \mathbf{Z}[\![x]\!]$$

is an algebraic function on $\mathbf{P}^1 \smallsetminus \{0, 1/16, \infty\}$, and the space of such $g(x)$ is infinite dimensional over $\mathbf{Q}(x)$ (by taking larger and larger N). Now let

$$h(x) = {}_3F_2\left[\begin{matrix} 1 & 1 & 1 \\ 1/2 & 1/2 \end{matrix}; \frac{x}{16}\right] = \sum \frac{x^n}{\binom{2n}{n}^2}.$$

The function $h(x)$ has denominator type $\tau = [1,2,3,\ldots,2n]^2$ and is a G-function of geometric origin over $\mathbf{P}^1 \smallsetminus \{0,16,\infty\}$. Now the Hadamard products

$$f(x) = g(x) \star h(x)$$

lie in $\mathcal{H}$ and also generate an infinite dimensional space over $\mathbf{Q}(x)$. $\square$

Note that the numerology in this case corresponds to $e^4 > 16$.

15.5. **Algorithmic Questions.** Given an ODE of geometric origin, one can generally always give a bound on the denominator type. However, determining the precise growth of the denominators appears to be difficult in general. Two enlightening examples can be given as follows. In [Coo12], the following example is considered. Let

$$x = q \prod_{n=1}^{\infty} \left(\frac{(1-q^{7n})}{(1-q^n)} \right)^4,$$

which is a Hauptmodul for $X_0(7)$. There is a corresponding uniformizer for $X_0(7)^+$ as follows:

$$y = \frac{x}{1 + 13x + 49x^2}.$$

If we now take the weight 2 Eisenstein series

$$E = \frac{7E_2(\tau) - E_2(7\tau)}{1-7} = 1 + 4q + 12q^2 + 16q^3 + 28q^4 + \ldots$$

and then write it in terms of y, we get an order three ODE without singularities on $\mathbf{P}^1 \smallsetminus \{0, 1/27, -1, \infty\}$, given explicitly by $\mathcal{L}H_A(x) = 0$ with

$$\mathcal{L} = x^2(1+x)(-1+27x)\frac{d^3}{dx^3} + 3x(-1+39x+54x^2)\frac{d^2}{dx^2} + (-1+86x+186x^2)\frac{d}{dx} + 4(1+6x).$$

If one considers the non-homogenous version $\mathcal{L}H(x) = -1$, then there exists a holomorphic solution $H(x)$ which is overconvergent beyond the cusp $1/27$. This solution is of modular origin, namely, there exists a weight 4 modular form whose triple (Eichler) integral gives rise to $H(x)/H_A(x)$ written in terms of the parameter q. What is unusual about this example is that the weight four form is meromorphic rather than holomorphic; it has the form h/E for the unique Hecke eigenform $h \in S_6(\Gamma_0(7), \mathbf{Q})$, and so h/E has poles away from the cusps. But perhaps more surprisingly, the form h/E is *magnetic* in the sense of [BZ19]; that is, if $h/E = \sum a_n q^n$ then $n | a_n$ for all n. As a consequence, the holomorphic solution $H(x)$, which from all appearances (and in light of the three integrations required to construct $H(x)$) one should expect to have denominator type $\tau = [1, 2, \ldots, n]^3$, actually has denominator type $\tau = [1, 2, \ldots, n]^2$. This is not at all apparent from the ODE, and it is not clear whether there is an algorithm to compute this *a priori*. As a curious consequence, it also means that the irrationality of the Apéry limit associated to Cooper's sequence is amenable to our methods; however, since the corresponding constant appears to be $\pi^2/42$, we have not pursued this!

The second example which highlights the difficulty in computing denominator types is as follows. Associated to Ramanujan's modular form $\Delta = \sum \tau(n) q^n$, one can write down an ODE with a non-homogeneous solution corresponding to the Eichler integral $\sum \tau(n) n^{-11} q^n$. One expects that the denominator type of the resulting function will be $[1, 2, \ldots, n]^{11}$. If one can prove that it is *not* of the form $A^n [1, 2, \ldots, n]^{10}$, however, then one would have proven that there exist infinitely many ordinary primes for Δ, a somewhat notorious open problem.

15.6. **The Gelfond–Schnirelman topic.** Recall the numerology of the very basic special case of Theorem 2.7.10 on which we based our arithmetic characterization of the logarithm: the slit plane domain $\Omega := \mathbf{C} \setminus [1, \infty)$ has conformal mapping radius $\rho(\Omega, 0) = 4$ with Riemann map $\varphi(z) = 4z/(1+z)^2$, and it admits the transcendental analytic function $\log(1-x) = -\sum_{n=1}^{\infty} x^n/n$ whose denominator type can be expressed into the form (2.7.8) with $r = 1$ and $b_1 = 1$. We have $(2/3)\log 4 = 0.924196\ldots$ for the right-hand side of (2.7.9) in this example.

This broaches another popular topic that was considered [Chu83a, § II], [FW08, pages 493–494], presumably for its methodological relevance, by several of the creators of the arithmetic theory of G-functions that we described in § 15.1. This is the old idea of Gelfond and Schnirelman who observed in 1936 that some prime counting lower bound $\pi(X) > (\log 2) X / \log X$ for all $X \gg 1$ follows at one stroke just by remarking upon the pointwise $\leq 4^{-n}$ integrand in $[1, \ldots, 2n+1] \int_0^1 (t - t^2)^n \, dt \in \mathbf{N}_{>0}$. Using the functional bad approximability property § 3.3.7 of $\log(1-x)$ (the normality of the Hermite–Padé table), the use of the prime number theorem in the proof of Theorem 2.7.2 can be turned around to devise a Gelfond–Schnirelman style elementary proof of the integrated prime counting function estimate $\int_1^X \psi(t)\, dt > (4/3) \log 2 \cdot X^2/2$, for all sufficiently big X. The coefficient here is slightly better than Chebyshev's $\log\left(2^{1/2} 3^{1/3} 5^{1/5} 30^{-1/30}\right) = 0.921292\ldots$, and now the point is that this proof is not really new: it is an isomorphic argument to Bombieri, Nair, and Chudnovsky's opening estimates which they got from using the discriminant[49] polynomial in the multivariable method; see [Chu83a, page 94], and note that the Cauchy determinant from the proof of [Nai82, Theorem 1] is none other than the Hankel determinant of the function $\log(1-x)$ as applied to our Remark 1.2.5. As the logic of the prime number theorem can be reversed in every single one of our arithmetic holonomy bounds, one cannot help but be curious about the denominator arithmetic in "asymptotic near-misses" of our holonomy bounds.

15.7. **A historical note and acknowledgments.** We originally conceived of our new approach to irrationality in 2020, starting with an easy proof of the irrationality of the 2-adic avatar of $\zeta(5)$ (which now we finally exposit in a companion paper [CDT24]), and even during that year we realized with the help of [Zag09] that the method could apply in principle to $L(2,\chi_{-3})$. However, at the time, the holonomy bounds that we could prove following [And89, § VIII] were totally insufficient, as explained in § 2. Our first serious improvements — such as (2.2.3) — over André's holonomy bound $(\sup_{\mathbf{T}} \log|\varphi|)\,/\,(\log|\varphi'(0)| - \tau)$ were still insufficient for this rather (as it seemed back then!) elusive application, but we found them nonetheless to carry a certain asymptotic precision which was the key to the proof [CDT21] of the unbounded denominators conjecture (the case $\tau = 0$ of algebraic functions). The paper [BC22] cites [CDT21] as a significant influence. In turn, [BC22], which implicitly already has the bound (2.2.5), has clearly been a crucial inspiration for our present paper, and (in part) it was by trying to synthesize our ideas with those of [BC22] that lead to the optimal holonomy bounds here.

The whole § 8.1 is due to Fedja Nazarov. We are grateful to him for explaining to us the precise analytic comparison between the Bost–Charles integral and the rearrangement integral. Remark 6.0.16 is based on a discussion with Samuel Goodman.

In addition, we would like to thank a number of people for conversations throughout the past four years on ideas related to this paper, including Yves André, Jean-Benoît Bost, Alin Bostan, François Charles, David and Gregory Chudnovsky, Tom Hutchcroft, Javier Frésan, Lars Kühne, Peter Sarnak, Umberto Zannier, Wadim Zudilin.

Appendix A. Choosing a contour

Recall (9.0.1) the function h defined as follows:

$$h := \lambda + \frac{\lambda}{\lambda - 1} = -256q\prod_{n=1}^{\infty}(1+q^n)^{24} = -256\cdot\frac{\Delta(2\tau)}{\Delta(\tau)}, \quad q = e^{2\pi i\tau}. \tag{A.0.1}$$

For any biholomorphic map $\psi : \mathbf{D} \to \Omega \subset \mathbf{D}$ with $\psi(0) = 0$, let $\varphi = h(\psi(x))$. Our task is to choose a function ψ for which:

(1) The image of ψ inside $\mathbf{D}$ avoids all preimages of the point $-1/72$ under h except for the one preimage $0.0000541\ldots \in \mathbf{R}$. (This is the only preimage on the real line.)

[49] In Bombieri's case, this inquiry led to the re-discovery of the Selberg integral and, bearing with this for the true mathematical fruit, the historic proof of the Dyson–Mehta conjecture along with cases of the Macdonald conjectures. This is the story recounted in [FW08, §§ 1.2, 1.3].

(2) The quantity (compare equation (7.0.3))

$$\frac{\iint_{\mathbf{T}^2} \log|\varphi(z)-\varphi(w)|\,\mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w)}{\log|256\psi'(0)|-\tau(\mathbf{b};\mathbf{e})} \tag{A.0.2}$$

is as small as possible, for a certain explicit constant

$$\tau(\mathbf{b};\mathbf{e}) = \frac{16603}{3920} = \frac{27}{80} + \frac{191}{49}.$$

Here $\log|\psi'(0)|$ is the conformal radius of Ω. The friction here is that we want the denominator of (A.0.2) to be large, and so $\Omega \subset \mathbf{D}$ to be large; at the same time, the function $h(z)$ has asymptotic growth

$$\log|h(z)| \sim \frac{1}{2\pi^2 q^2(1-|z|)}$$

as z varies in a straight line from 0 to the cusp $e^{2\pi i\alpha}$ where $\alpha = p/q$ and q is odd. (All of this follows easily from the fact that h is a modular function of level $\Gamma_0(2)$.) In order to choose Ω, it is instructive first to examine the topography of h. The shaded region in Figure A.0.3 indicates the $|z| \le 1$ for which $|h(z)| \ge e^{20}$, the level sets $|h(z)| \in \{1, e^4, e^8, e^{12}, e^{16}\}$, and then finally level sets $|h(z)+1/72| = 1/200$ around the preimages of $-1/72$. (The two types of level sets can be distinguished by whether any connected component has a subsequence tending towards the boundary or not.) The basic idea for constructing a Ω is to choose a circle centered at the origin with radius avoiding the preimages of $-1/72$ in the vicinity of $z = \pm i$, and then to (approximately) remove from Ω the following:

(1) The intersection of this circle with a horoball near $z = 1$.
(2) Slits from this circle to the remaining preimages of $-1/72$ along the (approximate) horoball in the vicinity of $z = -1$.

The Riemann mapping theorem guarantees the existence of a $\psi(x)$ for any such region Ω. However, we additionally want to choose $\psi(x)$ in an explicit form as follows in order to be able to rigorously estimate (A.0.2). Thus in practice we choose simple explicit functions which approximate this region. The construction of Ω is very much bespoke, and it is completely unclear (to us!) how to actually minimize (A.0.2) over all Ω, except to say from our experience that we believe our construction is not a long way from being optimal. Our ultimate choice of Ω is displayed in Figure A.4.5 (which also has a more detailed topographic map of $\log|h(z)|$).

A.1. **Preliminaries on Lunes.** Let $D(c,R)$ denote the disc of radius R centered at c. Fix $c > 1$, and consider the map

$$f(z,c) = z \cdot \frac{(c^2+1)+(c^2-1)z}{(c^2-1)+(c^2+1)z}.$$

This map has the following property; it is a conformal map from the lune $L(c)$ consisting of

$$L(c) := \mathbf{D} \smallsetminus \mathbf{D} \cap D\left(-\frac{c^2+1}{c^2-1}, \frac{2c}{c^2-1}\right)$$

to the unit disc sending $z = 0$ to 0. That is, the unit disc minus the intersection of two discs which intersect at $|z| = 1$ at right angles. (This guarantees the existence

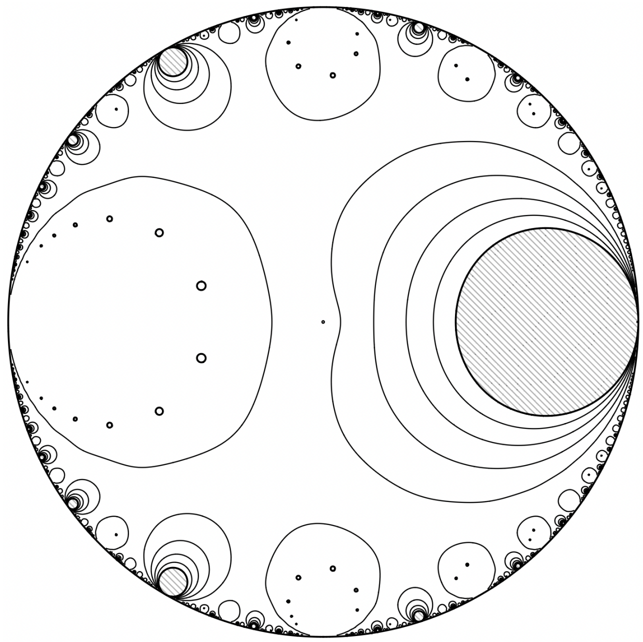

FIGURE A.0.3. The shaded region consists of $|z| \le 1$ with $|h(z)| \ge e^{20}$ near the cusps $z = e^{2\pi i p/q}$ with q odd. The other curves are the level sets $|h(z)| \in \{1, e^4, e^8, e^{12} e^{16}\}$, as well as the level sets $|h(z) + 1/72| = 1/200$ around the preimages of $-1/72$. The connected components of the former level sets are distinguished from the latter by containing subsequences converging to the boundary.

of an explicit and elementary conformal map). The "innermost" point of the right circle is the point

$$-\frac{c^2+1}{c^2-1} + \frac{2c}{c^2-1} = -\frac{c-1}{c+1}.$$

In the limit $c \to \infty$, this point tends to -1, and the region $L(c)$ tends to the entire disc, and $f(z)$ tends to z. The function $f(z)$ has an explicit inverse map as follows:

$$h(z,c) = \frac{z(1+c^2) - 1 - c^2 + \sqrt{(1+c^2)^2(1+z)^2 - 16c^2 z}}{2(c^2-1)}, \tag{A.1.1}$$

and hence the conformal radius of $L(c)$ is

$$\frac{c^2-1}{c^2+1}. \tag{A.1.2}$$

A.2. **Gobbles.** We do not use the contours of this section in the proofs of Theorems A or C, having replaced them by a combination of lunes with the slits considered in § A.3 below. However, they are used in the proof of Theorem 2.8.4 in § 6.8. Moreover, preliminary versions of our argument did employ them, and they do provide convenient contours on which to provide benchmarks for other examples, and are also more flexible than our somewhat custom use of slits.

Suppose we wish to remove *two* symmetrically opposite discs, not necessarily of equal sizes. There is no easily expressible simple conformal map in this case. However, as a first approximation, we can first remove one disc, and then remove the other. We define

$$\mathrm{Gob}(z,e,f) := h(-h(z,f),e).$$

Using equation (A.1.1), it has a somewhat messy but completely explicit form.

Definition A.2.1. The *gobble* inside $\mathbf{D}$ with parameters $r \in (0,1]$, $e \in (1,\infty]$, and $f \in (1,\infty]$ is the image of $\mathbf{D} = D(0,1)$ under the map

$$\mathrm{Gob}(r,e,f) : \mathbf{D} \to \mathbf{C},\ z \mapsto r \cdot \mathrm{Gob}(z,e,f).$$

One key property is that, for a wide range of parameters, the gobble is visually indistinguishable from the complement in $D(0,1)$ of two discs, while at the same time being much more explicit and thus easier to compute with. From the explicit formula, we easily obtain:

Lemma A.2.2. *The conformal radius of* $\mathrm{Gob}(r,e,f)(\mathbf{D})$ *centered at* $z=0$ *is equal to*

$$r \cdot \frac{(e^2-1)(f^2-1)}{(e^2+1)(f^2+1)}.$$

A.3. **Slits.** For a real number $r \in (0,1)$, a conformal isomorphism

$$(\mathbf{D},0) \xrightarrow{\simeq} (\mathbf{D} \smallsetminus (-1,-r],0)$$

is given by the function $\mathrm{Slit}(z,r)$ defined by the following formula:

$$\frac{(r+1)^2 - 2(r-1)^2 z + (r+1)^2 z^2 + (1+r)(-1+z)\sqrt{(1+r)^2 - 2(1-6r+r^2)z + (1+r)^2 z^2}}{8rz}$$
$$= \frac{4r}{(1+r)^2} z + \frac{8r(1-r)^2}{(1+r)^4} z^2 + \frac{4(1-r)^2 r(3-14r+3r^2)}{(1+r)^6} z^3 + \ldots . \tag{A.3.1}$$

In particular, the conformal radius of $\mathbf{D} \smallsetminus (-1,-r]$ at the origin is equal to

$$|\mathrm{Slit}'(0,r)| = \frac{4r}{(1+r)^2}. \tag{A.3.2}$$

We include a sketch of the derivation of the map (A.3.1). The starting point is to remark that the rational transformation

$$z \mapsto z/(1+z)^2,$$
$$0 \mapsto 0; \quad -1 \mapsto \infty; \quad 1 \mapsto 1/4; \quad -r \mapsto -r/(1-r)^2,$$
$$\text{with inverse } z \mapsto \frac{1-2z+\sqrt{1-4z}}{2z},$$

takes our slit disc $\mathbf{D} \smallsetminus (-1,-r]$ conformally isomorphically onto the $z \mapsto 1/z$ image of $\mathbf{P}^1 \smallsetminus \left[-(1-r)^2/r, 4\right]$. Now a line segment $[A,B] \subset \mathbf{R}$ has transfinite diameter ($[\infty]$-*capacitance*) equal to a quarter of its length. This already proves the formula (A.3.2) on the conformal size, for the complement in $\mathbf{P}^1 = \mathbf{C} \cup \{\infty\}$ of a contractible compact $K \subset \mathbf{C}$ is a topological disc whose Riemann mapping radius from ∞ is equal to the reciprocal of the transfinite diameter of K. For the actual Riemann map (A.3.1), we continue further by observing that the inverse Riemann map from ∞ for a segment complement $\mathbf{P}^1 \smallsetminus [A,B]$ in the Riemann sphere is given by a square root function:

$$\mathbf{P}^1 \smallsetminus [A,B] \xrightarrow{\simeq} \mathbf{H}, \quad z \mapsto i\sqrt{(z-A)/(z-B)}, \quad \infty \mapsto i,$$
$$\text{with inverse } z \mapsto \frac{A+Bz^2}{1+z^2}, \quad i \mapsto \infty.$$

Following this through by the Cayley transform $z \mapsto (z-i)/(z+i)$, $\mathbf{H} \xrightarrow{\simeq} \mathbf{D}$, whose inversion is $z \mapsto i\,(1+z)/(1-z)$, we arrive at the composed Riemann map (A.3.1).

A.4. **Combining multiple slits and lunes.** Suppose we wish to remove *four* slits and, additionally, a lune. There is no easily expressible simple conformal map in this case. However, both the conformal maps in (A.1.1) and (A.3.1) have the property that they are well–approximated by the identity map $z \mapsto z$ for "most" points in the circle (namely, the points away from the lune and the slits, respectively). Thus one very primitive way to construct such maps is simply to *compose* these maps in succession. To this end, we consider the following function:

$$G(z) := -R{\cdot}h\left(-e^{2\pi i\theta_1}\cdot \mathrm{Slit}\left(e^{2\pi i\theta_2}\cdot \mathrm{Slit}\left(e^{2\pi i\theta_3}\mathrm{Slit}\left(e^{2\pi i\theta_4}\cdot \mathrm{Slit}\left(z,r_1\right),r_2\right),r_3\right),r_4\right),c\right).$$

We fix the first parameter $R = 77/100$ to ensure that the initial circle only contains preimages of $-1/72$ in the horoball around -1, and indeed that there are only 4 such preimages that we need to exclude. We also fix the lune parameter $c = 75/10$ which measures the (approximate) horoball we remove near $z = 1$. The angle parameters θ_i allow us to "line up" the slits so that they include these preimages, and the length parameters r_i allow us to minimize the lengths of these slits so they do not go beyond the preimages we wish to exclude. The final choice of parameters is as follows:

$$\begin{aligned} &R = \frac{77}{100}, \qquad c = \frac{75}{10},\\ &r_1 = \frac{91}{100}, \quad r_2 = \frac{6188}{10000}, \quad r_3 = \frac{55515}{100000}, \quad r_4 = \frac{772}{1000},\\ &\theta_1 = \frac{7977}{100000}, \quad \theta_2 = \frac{11543}{100000}, \quad \theta_3 = \frac{3525}{100000}, \quad \theta_4 = -\frac{783}{10000}. \end{aligned} \tag{A.4.1}$$

These parameters are chosen from an *ad hoc* computation making the ends of the slits as close to the four parameters as possible. Since it is not possible (numerically) to choose these parameters so that the bad preimages lie exactly on these slits, we finally define

$$\psi(z) = G\left(\frac{995}{1000}\cdot z\right). \tag{A.4.2}$$

By restricting to this open disc, we are removing not simply (curved) slits but open regions, which enables one to easily prove that the bad preimages are excluded. It is simple enough to compute that the conformal radius of $\psi(z)$ (using (A.1.2) and (A.3.2)) and it is equal to

$$\begin{aligned} |\psi'(0)| &= \frac{995}{1000}\cdot R\cdot \frac{c^2-1}{c^2+1}\cdot \prod_{i=1}^{4}\frac{4r_i}{(1+r_i)^2}\\ &= \frac{5448339453535586608000000000}{8658833407565631122430056127} = 0.6292232680\ldots \end{aligned} \tag{A.4.3}$$

Recall that $h : \mathbf{D} \to \mathbf{C}$ is given by $-256q\prod_{n=1}^{\infty}(1+q^n)^{24}$ with $q \in \mathbf{D}$. The parameters r_i and θ_i are chosen above to ensure that the following holds:

Lemma A.4.4. *The map $\varphi := h \circ \psi : \overline{\mathbf{D}} \to \mathbf{C}$ has a unique preimage of $-1/72$.*

Proof. There is a preimage with $x = 0.0000541829\ldots$. But one can determine (to any precision) the other preimages by passing back to $\mathbf{H}$ and then the preimages are obtained by the action of $\Gamma_0(2)$. The preimages in the region (approximating a horoball) near $z = \pm i$ have absolute value at least

$$0.782767\ldots > R = \frac{77}{100}.$$

The closest other preimages lie near the horoball at $z = -1$; but the precise choice of the parameters r_i and θ_i ensure that they lie outside image of ψ as can be confirmed by a simple numerical computation. □

The contour $\psi(\mathbf{T})$ is drawn in Figure A.4.5. The asymmetry is due to our using four successive compositions of single slit maps, rather than having all four slits taken out at once.

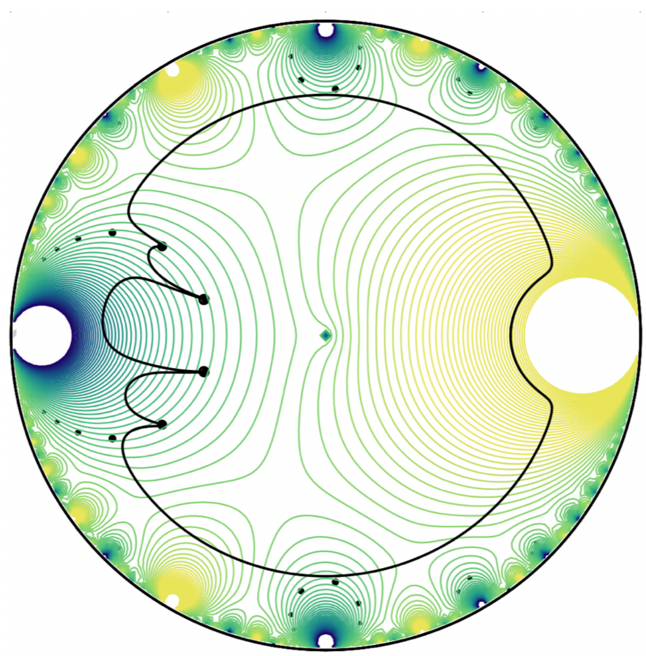

FIGURE A.4.5. The image of $|z| = 1$ under $\psi(z)$ together with preimages of $-1/72$ under $h : \mathbf{D} \to \mathbf{C}$, together with level sets for $\log|h|$ at values in an arithmetic progression; the color scheme transitions between yellow for large positive values of $\log|h(z)|$ and blue for large negative values.

Remark A.4.6. Our choice of constants reflects merely the principle of finding an example "which works" rather than is the most aesthetically pleasing. There is no doubt some scope for improvement but since it is not necessary we have not tried to optimize these choices — we expect improvements in either respect would anyway be quite modest. △

A.5. **A contour for the $L(2,\chi_{-3})$ problem.** With our specific choice of $\psi(z)$ as in Definition A.4.2, we now define

$$\varphi(z) = h(\psi(z)),$$

which is uniformly continuous on $\mathbf{D}$, and which is explicit enough as to be amenable to rigorous numerical estimates. We then finally obtain the estimate

$$\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\, \mu_{\mathrm{Haar}}(z)\mu_{\mathrm{Haar}}(w) = 11.844\ldots$$

and thus (A.0.2) is bounded above by

$$\frac{11.845}{\log\left(256 \cdot \frac{544833945353558660800000000}{865883340756563112243005612 7}\right) - \left(\frac{27}{80} + \frac{191}{49}\right)} = 13.9938\ldots. < 14. \tag{A.5.1}$$

Remark A.5.2 (Bounds without integrations)**.** Even with all our improvements, the best bound we could achieve before integrations, for either Theorems A or C, was also above 9; we give some of the numerics now. Consider the following basic application of Theorem 2.5.1. We consider the functions $A_i(x)$ for $i = 1,\ldots,9$ in the $\mathbf{P}^1 \smallsetminus \{0,1,\infty\}$ domain. (Here the first five functions are given explicitly in § 10, and the four functions $A_6(x),\ldots,A_9(x)$ correspond to $B_6(z),\ldots,B_9(z)$ via the transformations of that section. Note that these last four functions only exist if there is a $\mathbf{Q}$-linear relation between our three periods.) Consider Theorem 2.5.1 for

$$\mathbf{b} := \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{t}},$$

hence

$$\sigma_1 = 0,\ \sigma_2 = 1; \quad \sigma_3 = \cdots = \sigma_9 = 2,$$

and

$$\tau(\mathbf{b}) = \frac{1\cdot 0 + 3\cdot 1 + (5+7+9+11+13+15+17)\cdot 2}{81} = \frac{157}{81}.$$

As explained in Remark 9.0.20, if we use the same contour as given in Definition A.4.2 except pulled back to the $X(2)$ domain, both the integral and the conformal radius terms are doubled. Equivalently, they remain the same and the τ term is halved. Hence the corresponding bound we obtain in this case is:

$$\frac{11.844\ldots}{5.081\ldots - 2\cdot 157/81} = 9.833\ldots < 10. \tag{A.5.3}$$

which comes close but is not a contradiction because this term is not less than 9. While this can be refined slightly (using the Bost–Charles integral and modifying the contour), it seems unlikely that one may reach a direct contradiction by our methods without involving added integrations; see Examples 7.4.9 and 7.5.9. △

A.6. **A contour for the logarithm problem.** We could literally use the same contour as above to complete the proof of Theorem C, except with a somewhat worse constant. Following the arguments of Section 14.4, it would suffice to find the ε such that the image of φ above excludes the regions where z is not too small and $h(z)$ lies in $D(0,\varepsilon^2/16)$ and also where $h(z)$ lies outside $D(0,\varepsilon^{-1})$. This leads to a choice of ε somewhere between 10^7 and 10^8. However, a compromise between optimizing over various conformal maps and using the same map as above is just to write down simple lunes: If we take

$$\psi(z) = -\frac{3}{4} h\left(z, \frac{23}{10}\right)$$

of conformal radius 1287/2516, then the image of $\psi(z)$ avoids all the required discs as well as the regions where $|h(z)| \geq 10^6$ and $|h(z)| \leq 10^{-12}2^{-4}$ (except for the preimages near $z = 0$). In this case, we obtain the bound

$$\iint_{\mathbf{T}^2} \log|\varphi(z) - \varphi(w)|\, \mu_{\text{Haar}}(z)\mu_{\text{Haar}}(w) \sim 9.963\ldots < 10, \tag{A.6.1}$$

and we have the very easy bound

$$\frac{10}{\log\left(256 \cdot \frac{1287}{2516}\right) - \frac{1032659}{242760}} = 16.103\ldots < 17 \tag{A.6.2}$$

which we use for the proof of Theorem C in § 14.5. In comparison, we may also estimate the rearrangement integral

$$\int_0^1 2t \cdot (\log|\varphi(e^{2\pi i t})|)^* \, dt \sim 9.972\ldots \tag{A.6.3}$$

which also suffices to prove Theorem C, this time via Theorem 6.0.2 with only the trivial partition of $[0, m]$.

A graph of the image of $\psi(z)$ together with the regions where $|h| \geq 10^6$ and $|h| \leq 10^{-12}2^{-4}$ is given in Figure A.6.4

Appendix B. A dynamic box principle

In this appendix we give a short new proof of the basic holonomy bound

$$m \leq \frac{2T(\varphi)}{\log|\varphi'(0)| - b_1 - \cdots - b_r}, \tag{B.0.1}$$

under the condition of the positive denominator, for a $\mathbf{Q}(x)$-linearly independent set of formal functions $f_1, \ldots, f_m \in \mathbf{Q}[\![x]\!]$ of the types (2.6.2) and such that $\varphi; \varphi^* f_1, \ldots, \varphi^* f_m \in \mathcal{M}(\overline{\mathbf{D}})$ are simultaneously meromorphic on a neighborhood of the closed unit disc $\overline{\mathbf{D}}$. Here,

$$T(\varphi) := \int_{\mathbf{T}} \log^+|\varphi|\, \mu_{\text{Haar}} + \sum_{\substack{\rho \in \mathbf{D} \\ \text{poles of } \varphi}} \log \frac{1}{\rho} \tag{B.0.2}$$

is the *Nevanlinna characteristic* of the meromorphic mapping φ (the meromorphic poles being taken with their multiplicities).

This is based on the idea of Perelli and Zannier [PZ84] with a dynamic box principle such as they formulate with their Lemma 1 of *loc.cit.* It may be considered as a more elementary form of Bost's technique in § 7, to which it is both an introduction and an alternative, and to our companion paper [CDT24], where these ideas are pursued further. We divide the proof into three steps according to the dissection in § 2.12.

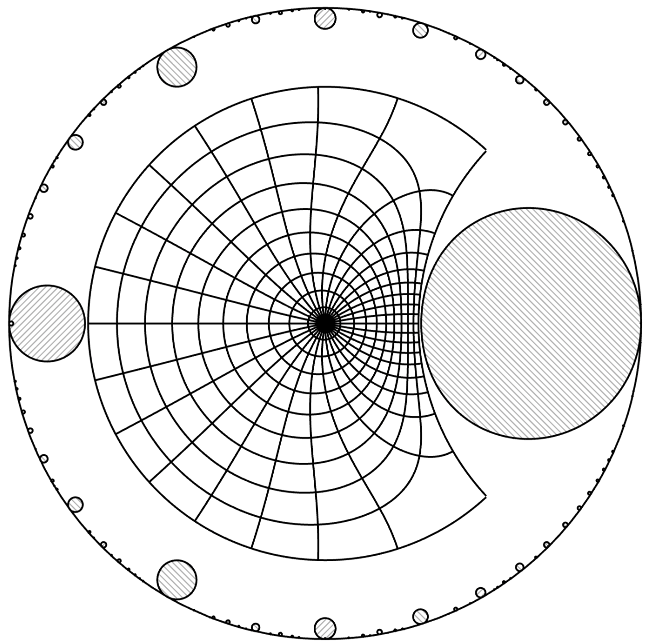

FIGURE A.6.4. The image of $|z| = 1$ together with the images of the level sets $|z| = k/10$ and $\arg(z) = 2\pi k/32$ for integers k, together with the regions with $|h| \geq 10^6$ and $|h| < 10^{-12}2^{-4}$ (distinguished by the angle of the shading)

B.1. **Evaluation module.** Suppose we have a $\mathbf{Q}(x)$-linearly independent set of functions $f_1(x), \ldots, f_m(x) \in \mathbf{Q}[\![x]\!]$ of the type (2.6.2) such that $\varphi(z) \in \mathbf{C}[\![z]\!]$, as well as each power series $f_i(\varphi(z)) \in \mathbf{C}[\![z]\!]$, are germs of meromorphic functions on a neighborhood of the closed unit complex disc $|z| \leq 1$. (Having this slightly bigger disc is no loss of generality upon replacing $\varphi(z)$ by $\varphi(\rho z)$ for some $\rho < 1$ still having $\rho|\varphi'(0)| > e^{b_1+\cdots+b_r}$.) We introduce two positive integer parameters D and T, and we consider the collection

$$\mathcal{I}_D(T) := \left\{ (Q_1, \ldots, Q_m) \in \mathbf{Z}[x] \,:\, Q_i(x) = \sum_{j=0}^{D-1} c_{i,j} x^j,\ c_{i,j} \in [0, T) \cap \mathbf{Z} \right\}, \tag{B.1.1}$$

of cardinality

$$\#\mathcal{I}_D = T^{mD}. \tag{B.1.2}$$

By the assumed $\mathbf{Q}(x)$-linear independence of the m formal power series $f_i(x) \in \mathbf{Q}[\![x]\!]$, the $\mathbf{Z}$-module *evaluation map*

$$\psi_D \,:\, \mathbf{Z}[x]^{\oplus m}_{\deg < D} \hookrightarrow \mathbf{Q}[\![x]\!], \qquad (Q_1, \ldots, Q_m) \mapsto \sum_{i=1}^{m} Q_i(x) f_i(x) \in \mathbf{Q}[\![x]\!] \tag{B.1.3}$$

is *injective*. Hence, the image

$$\mathcal{O}_D := \psi_D(\mathcal{I}_D) \subset \mathbf{Q}[\![x]\!]$$

under this map also has cardinality

$$\#\mathcal{O}_D = \#\mathcal{I}_D = T^{mD}. \tag{B.1.4}$$

In the free **Z**-module (B.1.3) of rank mD, we define the *vanishing filtration jumps*

$$r_D^{(n)} := \dim_{\mathbf{Q}} \left\{ \left(\psi_D^{-1}\left(x^n \mathbf{Q}[\![x]\!]\right) \otimes \mathbf{Q}\right) / \left(\psi_D^{-1}\left(x^{n+1}\mathbf{Q}[\![x]\!]\right) \otimes \mathbf{Q}\right) \right\}. \tag{B.1.5}$$

They are in $\{0,1\}$, because the linear injective map ψ_D induces a linear injection

$$\left(\psi_D^{-1}\left(x^n \mathbf{Q}[\![x]\!]\right) \otimes \mathbf{Q}\right) / \left(\psi_D^{-1}\left(x^{n+1}\mathbf{Q}[\![x]\!]\right) \otimes \mathbf{Q}\right) \hookrightarrow x^n\mathbf{Q}[\![x]\!]/x^{n+1}\mathbf{Q}[\![x]\!] \cong \mathbf{Q}\cdot x^n$$

into a one-dimensional **Q**-vector space. On the other hand, we have

$$\sum_{n=0}^{\infty} r_D^{(n)} = \dim_{\mathbf{Q}} \psi_D^{-1}\mathbf{Q}[\![x]\!] = mD. \tag{B.1.6}$$

Hence there is a size-mD set of possible $x=0$ vanishing orders

$$\begin{aligned} &\left\{ n \in \mathbf{N} \,:\, \exists (Q_1,\dots,Q_m) \in \mathbf{Z}[x]^{\oplus m}_{\deg < D},\ \operatorname{ord}_{x=0}\left(\sum_{i=1}^m Q_i(x)f_i(x)\right) = n \right\} \\ &= \left\{ 0 \le u(1) < u(2) < \cdots < u(mD) \right\} \end{aligned} \tag{B.1.7}$$

for our auxiliary functions. They depend only on the module (B.1.3) — in other words, on $f_1,\dots,f_m$ and the parameter D, — but not on the parameter T, which remains free to select in the following. (The parameter T will be taken to be any sufficiently big integer in dependence of the filtration jumps (B.1.7).)

This fulfills step (i) of § 2.12.

B.2. **Box principle.** For step (ii), we measure up the tendency of the Taylor series of the auxiliary function $F(x)$ to depend recursively on its string of initial coefficients under the critical condition $|\varphi'(0)| > e^{b_1+\ldots+b_r}$.

We can upper-estimate the output cardinality $\#\mathcal{O}_D$ by a product $\prod_{p=1}^{mD}\gamma_p$, where γ_p is an upper estimate on the largest possible number of distinct $x^{u(p)}$ coefficients $\beta_{u(p)} \in \mathbf{Q}$ in any set of output functions $F(x) = \sum_{k=0}^{\infty}\beta_k x^k$ that share a common string $(\beta_0, \beta_1, \dots, \beta_{u(p)-1})$ for their preceding coefficients:

$$\begin{aligned} &\forall (\beta_0,\dots,\beta_{u(p)-1}) \in \mathbf{Q}^{u(p)}, \\ &\#\left\{ \beta \in \mathbf{Q} \,:\, \exists (Q_1,\dots,Q_m) \in \mathcal{I}_D,\ \sum_{i=1}^m Q_i(x)f_i(x) - \sum_{k=0}^{u(p)-1}\beta_k x^k \right. \\ &\qquad \left. = \beta x^{u(p)} + O\big(x^{u(p)+1}\big) \right\} \le \gamma_{u(p)} \end{aligned}$$

At this point, the integrality condition (2.6.2) is used to remark that all such rational numbers $\beta \in \mathbf{Q}$ belong in fact to a **Z**-module given by the requisite denominators type:

$$\beta \in \frac{1}{[1,\dots,b_1u(p)]\cdots[1,\dots,b_ru(p)]}\,\mathbf{Z}.$$

Hence, if $A_p \in \mathbf{R}^{>0}$ is such that any two such coefficients β differ by some real number in $[-A_p, A_p]$, then we can take

$$\gamma_p := 1 + 2A_p\cdot[1,\dots,b_1u(p)]\cdots[1,\dots,b_ru(p)] = 1 + A_p \cdot e^{(b_1+\ldots+b_r)u(p)+o(u(p))+O(1)} \tag{B.2.1}$$

as a total bound on the output possibilities of $\beta_{u(p)}$ given $(\beta_0, \ldots, \beta_{u(p)-1})$. This step emulates the dynamic box principle of Perelli and Zannier [PZ84, § 2 Lemma 1].

B.3. **Diophantine analysis of the lowest order coefficient.** The bound for A_p comes analytically from using the simultaneous meromorphic uniformization map φ. Consider $\varphi = v/u$ any representation of φ as the quotient of two convergent power series u and v on $\overline{\mathbf{D}}$ such that $u(0) = 1$. Let h be a convergent power series on $\overline{\mathbf{D}}$ such that $h(0) = 1$ and hf_i is holomorphic for each $i = 1, \ldots, m$. Let us write $n := u(p)$ for the following. Any two output functions $F_1(x)$ and $F_2(x)$ as above whose $x = 0$ Taylor series coincide up to $O(x^n)$, and whose respective x^n coefficients are $\beta^{(1)}$ and $\beta^{(2)}$, will have

$$V(z) := h(z)u(z)^D \cdot (F_1(\varphi(z)) - F_2(\varphi(z)))$$

holomorphic (convergent) on a neighborhood of the closed unit disc $\overline{\mathbf{D}}$, and with leading order term $\varphi'(0)^n(\beta^{(1)} - \beta^{(2)})z^n + O(z^{n+1})$ expressible as a $|z| = 1$ contour integral by Cauchy's formula:

$$\varphi'(0)^n(\beta^{(1)} - \beta^{(2)}) = \int_{\mathbf{T}} \frac{V(z)}{z^{n+1}}\, \mu_{\mathrm{Haar}}. \tag{B.3.1}$$

Estimating by the supremum of the integrand, we can take for our A_p the upper bound on the bottom row of

$$\begin{aligned} \left|\beta^{(1)} - \beta^{(2)}\right| &\leq |\varphi'(0)|^{-n} \cdot \sup_{\mathbf{T}} |V| \\ &\leq |\varphi'(0)|^{-u(p)} \cdot T \left(\sup_{\mathbf{T}} \max\left(|u|, |v|\right)\right)^D \cdot mD \cdot \sup_{\mathbf{T}} |h \cdot \varphi^* f_i| =: A_p, \end{aligned} \tag{B.3.2}$$

used with $n := u(p)$.

We get for the $\prod_{p=1}^{mD} \gamma_p$ output possibilities the upper estimate $T^{mD} = \#\mathcal{O}_D$

$$\leq \prod_{p=1}^{mD} \Bigg\{ 1 + T \exp\Bigg(- \big(\log|\varphi'(0)| - b_1 - \cdots - b_r + o(1)\big) u(p) + D \sup_{\mathbf{T}} \log\max(|u|, |v|) + \log D + O_{m,h}(1) \Bigg) \Bigg\}.$$

At this point, we look at the last inequality asymptotically in $T \to \infty$, or more concretely, we select a T so big that all mD factors of the product are ≥ 2. Using the trivial inequality $1 + x \leq 2x$ for $x \geq 1$ and canceling the common ensuing T^{mD} from both sides, we get (after taking the logarithm)

$$\begin{aligned} -\big(\log|\varphi'(0)| - b_1 - \cdots - b_r \big)(1 - o(1)) \sum_{p=1}^{mD} u(p) + mD^2 \sup_{\mathbf{T}} \log\max(|u|, |v|) + O(D \log D) \\ \geq -mD \log 2 - O_{m,h}(D). \end{aligned}$$

As $0 \leq u(p) < u(2) < \cdots < u(mD)$ are a strictly increasing sequence of nonnegative integers, we have

$$\sum_{p=1}^{mD} u(p) \geq \sum_{n=0}^{mD-1} n = \binom{mD}{2}. \tag{B.3.3}$$

We derive

$$(1-o(1))\binom{mD}{2}\big(\log|\varphi'(0)|-b_1-\cdots-b_r\big) \leq mD^2 \sup_{\mathbf{T}} \log\max(|u|,|v|)+O_{m,h,1}(D), \tag{B.3.4}$$

which in the $D \to \infty$ asymptotic filters down to the *arithmetic holonomy bound*

$$m \leq \frac{2\sup_{\mathbf{T}} \log\max(|u|,|v|)}{\log|\varphi'(0)|-b_1-\cdots-b_r}. \tag{B.3.5}$$

This is true for any meromorphic quotient representation $\varphi = v/u$ with $u(0) = 1$. A well-known lemma of Nevanlinna (cf. [Nev70, § VII.1.4] or [Gol69, § VII.5]), based on the canonical Blaschke products and the canonical decomposition $\log = \log^+ - \log^-$ in the Poisson–Jensen formula, constructs *on the open disc* $\mathbf{D}$ a quotient representation $\varphi = v/u$ with $u(0) = 1$ and with both $\sup_{\mathbf{D}} |u|$ and $\sup_{\mathbf{D}} |v|$ bounded by $\exp(T(\varphi))$. Dilating the radius a little bit, we get for any $\varepsilon > 0$ a quotient representation $\varphi = v/u$, now on some neighborhood of the closed disc $\overline{\mathbf{D}}$ as required in the above analysis, with $\sup_{\mathbf{T}} |u|$ and $\sup_{\mathbf{T}} |v|$ bounded by $\exp(T(\varphi) + \varepsilon)$. This concludes the proof of the bound (B.0.1). □

References

[AB95] Ahmad Abbes and Thierry Bouche, *Théorème de Hilbert-Samuel "arithmétique"*, Ann. Inst. Fourier (Grenoble) **45** (1995), no. 2, 375–401. 98

[AB97] Yves André and Francesco Baldassarri, *Geometric theory of G-functions*, Arithmetic geometry (Cortona, 1994), Sympos. Math., vol. XXXVII, Cambridge Univ. Press, Cambridge, 1997, pp. 1–22. 192

[Ada87] Colin C. Adams, *The noncompact hyperbolic 3-manifold of minimal volume*, Proc. Amer. Math. Soc. **100** (1987), no. 4, 601–606. 4

[Ami75] Yvette Amice, *Les nombres p-adiques*, Collection SUP: "Le Mathématicien", vol. 14, Presses Universitaires de France, Paris, 1975, Préface de Ch. Pisot. 7, 193

[And89] Yves André, *G-Functions and Geometry*, Aspects of Mathematics, no. E13, Friedr. Vieweg Sohn, Braunschweig, 1989. 7, 12, 17, 37, 52, 53, 189, 190, 199

[And96] ———, *G-fonctions et transcendance*, J. Reine Angew. Math. **476** (1996), 95–125. 12

[And00a] ———, *Séries Gevrey de type arithmétique. I. Théorèmes de pureté et de dualité*, Ann. of Math. (2) **151** (2000), no. 2, 705–740. 193, 194, 196

[And00b] ———, *Séries Gevrey de type arithmétique. II. Transcendance sans transcendance*, Ann. of Math. (2) **151** (2000), no. 2, 741–756. 193, 194, 195, 196

[And04] ———, *Sur la conjecture des p-courbures de Grothendieck–Katz et un problème de Dwork*, Geometric Aspects of Dwork Theory, vol. I, de Gruyter, Berlin, 2004, pp. 55–112. 7, 12, 52

[Ang1919] Aurel Angelesco, *Sur deux extensions des fractions continues algébriques*, Comptes Rendues Hebdomadaires des Séances de l'Académie des Sciences **18** (1919), 262–263. 51

[Ape79] Roger Apery, *Irrationalité de ζ2 et ζ3*, Astérisque (1979), no. 61, 11–13, Luminy Conference on Arithmetic. 3, 18

[AR79] Krishna Alladi and Michael L. Robinson, *On certain irrational values of the logarithm*, Number theory, Carbondale 1979 (Proc. Southern Illinois Conf., Southern Illinois Univ., Carbondale, Ill., 1979), Lecture Notes in Math., vol. 751, Springer, Berlin, 1979, pp. 1–9. 174

[AR80] ———, *Legendre polynomials and irrationality*, J. Reine Angew. Math. **318** (1980), 137–155. 57, 174

[AZ90] Gert Almkvist and Doron Zeilberger, *The method of differentiating under the integral sign*, J. Symbolic Comput. **10** (1990), no. 6, 571–591. 163

[Bak64] Alan Baker, *Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers*, Quart. J. Math. Oxford Ser. (2) **15** (1964), 375–383. 56

[Bak22] ______, *Transcendental number theory*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2022, With an introduction by David Masser, Reprint of the 1975 original. 174

[BB85] Daniel Bertrand and Frits Beukers, *Équations différentielles linéaires et majorations de multiplicités*, Ann. Sci. École Norm. Sup. (4) **18** (1985), no. 1, 181–192. 38, 52

[BB98] Jonathan M. Borwein and Peter B. Borwein, *Pi and the AGM*, Canadian Mathematical Society Series of Monographs and Advanced Texts, vol. 4, John Wiley & Sons, Inc., New York, 1998, A study in analytic number theory and computational complexity, Reprint of the 1987 original, A Wiley-Interscience Publication. 30

[BBR90] Frits Beukers, Jean-Paul Bézivin, and Philippe Robba, *An alternative proof of the Lindemann-Weierstrass theorem*, Amer. Math. Monthly **97** (1990), no. 3, 193–197. 193

[BC97a] Enrico Bombieri and Paula B. Cohen, *Effective Diophantine approximation on $\mathbf{G}_m$. II*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **24** (1997), no. 2, 205–225. 54

[BC97b] ______, *Siegel's lemma, Padé approximations and Jacobians*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **25** (1997), no. 1-2, 155–178, With an appendix by Umberto Zannier, Dedicated to Ennio De Giorgi. 35, 192

[BC22] Jean-Benoît Bost and François Charles, *Projective and formal-analytic arithmetic surfaces*, 2022, https://arxiv.org/abs/2206.14242v2. 10, 11, 13, 14, 15, 17, 31, 44, 45, 54, 68, 71, 89, 90, 91, 93, 97, 98, 99, 100, 107, 126, 127, 131, 134, 135, 193, 194, 199

[BCY04] Daniel Bertrand, Vladimir Chirskii, and Johan Yebbou, *Effective estimates for global relations on Euler-type series*, Ann. Fac. Sci. Toulouse Math. (6) **13** (2004), no. 2, 241–260. 38, 52

[Ber99] Daniel Bertrand, *On André's proof of the Siegel–Shidlovsky theorem*, Colloque Franco-Japonais: Théorie des Nombres Transcendants (Tokyo, 1998), Sem. Math. Sci., vol. 27, Keio Univ., Yokohama, 1999, pp. 51–63. 38

[Ber12] ______, *Le théorème de Siegel–Shidlovsky revisité*, Number theory, analysis and geometry, Springer, New York, 2012, pp. 51–67. 38, 52

[Beu79] Frits Beukers, *A note on the irrationality of $\zeta(2)$ and $\zeta(3)$*, Bull. London Math. Soc. **11** (1979), no. 3, 268–272. 162

[Beu81] ______, *Padé-approximations in number theory*, Padé approximation and its applications, Amsterdam 1980 (Amsterdam, 1980), Lecture Notes in Math., vol. 888, Springer, Berlin-New York, 1981, pp. 90–99. 192

[Beu84] ______, *The values of polylogarithms*, Topics in classical number theory, Vol. I, II (Budapest, 1981), Colloq. Math. Soc. János Bolyai, vol. 34, North-Holland, Amsterdam, 1984, pp. 219–228. 192

[Beu87] ______, *Irrationality proofs using modular forms*, Astérisque (1987), no. 147-148, 271–283, 345, Journées arithmétiques de Besançon (Besançon, 1985). 19, 159, 161, 162

[Beu06] ______, *A refined version of the Siegel–Shidlovskii theorem*, Ann. of Math. (2) **163** (2006), no. 1, 369–379. 193, 194, 195

[BG06] Enrico Bombieri and Walter Gubler, *Heights in Diophantine Geometry*, Cambridge New Mathematical Monographs, no. 4, Cambridge University Press, 2006. 7, 23, 35, 60, 75, 76, 78, 91, 182, 190

[BGMN05] Franck Barthe, Olivier Guédon, Shahar Mendelson, and Assaf Naor, *A probabilistic approach to the geometry of the l_p^n-ball*, Ann. Probab. **33** (2005), no. 2, 480–513. 59

[BGS94] Jean-Benoît Bost, Henri Gillet, and Christophe Soulé, *Heights of projective varieties and positive Green forms*, J. Amer. Math. Soc. **7** (1994), no. 4, 903–1027. 4, 137

[Bin16] Gal Binyamini, *Multiplicity estimates: a Morse-theoretic approach*, Duke Math. J. **165** (2016), no. 1, 95–128. 53

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux. 61

[BM80] W. Dale Brownawell and David W. Masser, *Multiplicity estimates for analytic functions. II*, Duke Math. J. **47** (1980), no. 2, 273–295. 53

[BM83] Enrico Bombieri and Julia Mueller, *On effective measures of irrationality for $\sqrt[r]{a/b}$ and related numbers*, J. Reine Angew. Math. **342** (1983), 173–196. 54

[Bom81] Enrico Bombieri, *On G-functions*, Recent progress in analytic number theory, Vol. 2 (Durham, 1979), Academic Press, London-New York, 1981, pp. 1–67. 57, 189, 190, 192

[Bom82] ———, *On the Thue–Siegel–Dyson theorem*, Acta Math. **148** (1982), 255–296. 54, 192

[Bom83] ———, *On Weil's "théorème de décomposition"*, Amer. J. Math. **105** (1983), no. 2, 295–308. 190

[Bom93] ———, *Effective Diophantine approximation on* $\mathbf{G}_m$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **20** (1993), no. 1, 61–89. 54

[Bor1914] Émile Borel, *Introduction géométrique à quelques théories physiques*, 1914. 59

[Bos99] Jean-Benoît Bost, *Potential theory and Lefschetz theorems for arithmetic surfaces*, Ann. Sci. École Norm. Sup. (4) **32** (1999), no. 2, 241–312. 98, 131

[Bos01] ———, *Algebraic leaves of algebraic foliations over number fields*, Publ. Math. Inst. Hautes Études Sci. (2001), no. 93, 161–221. 44, 94, 96, 97, 132, 133, 134

[Bos04] ———, *Germs of analytic varieties in algebraic varieties: canonical metrics and arithmetic algebraization theorems*, Geometric aspects of Dwork theory. Vol. I, Walter de Gruyter, Berlin, 2004, pp. 371–418. 44, 132

[Bos20] ———, *Theta invariants of Euclidean lattices and infinite-dimensional Hermitian vector bundles over arithmetic curves*, Progress in Mathematics, vol. 334, Birkhäuser/Springer, 2020. 13, 44, 45, 90, 94, 97, 98, 100, 105, 134, 136, 194

[Boy98] David W. Boyd, *Mahler's measure and special values of L-functions*, Experiment. Math. **7** (1998), no. 1, 37–82. 31

[BR89] Jean-Paul Bézivin and Philippe Robba, *A new p-adic method for proving irrationality and transcendence results*, Ann. of Math. (2) **129** (1989), no. 1, 151–160. 193, 195

[BR01] Keith Ball and Tanguy Rivoal, *Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs*, Invent. Math. **146** (2001), no. 1, 193–207. 5, 51

[BV89] Jean-Michel Bismut and Éric Vasserot, *The asymptotics of the Ray–Singer analytic torsion associated with high powers of a positive line bundle*, Comm. Math. Phys. **125** (1989), no. 2, 355–367. 98

[BZ19] David Broadhurst and Wadim Zudilin, *A magnetic double integral*, J. Aust. Math. Soc. **107** (2019), no. 1, 9–25. 198

[BZ20] François Brunault and Wadim Zudilin, *Many variations of Mahler measures—a lasting symphony*, Australian Mathematical Society Lecture Series, vol. 28, Cambridge University Press, Cambridge, 2020. 4

[BZ22] Francis Brown and Wadim Zudilin, *On cellular rational approximations to* $\zeta(5)$, 2022, https://arxiv.org/abs/2210.03391v2. 192

[Car54] Constantin Carathéodory, *Theory of Functions of a Complex Variable. Vol. 2*, Chelsea Publishing Co., New York, 1954, Translated by F. Steinhardt. 24, 152

[Cat1882] Eugène Charles Catalan, *Recherches sur la constante G, et sur les intégrales eulériennes*, Mémoires de l'Académie Impériale des Sciences de St. Pétersbourg **Tome XXXI** (1882), 55 p. 154

[CC83] David V. Chudnovsky and Gregory V. Chudnovsky, *Rational approximations to solutions of linear differential equations*, Proc. Nat. Acad. Sci. U.S.A. **80** (1983), no. 16, 5158–5162. 38, 53, 54

[CC85a] ———, *Applications of Padé approximations to Diophantine inequalities in values of G-functions*, Number theory (New York, 1983–84), Lecture Notes in Math., vol. 1135, Springer, Berlin, 1985, pp. 9–51. 52, 53, 57, 189, 190

[CC85b] ———, *Applications of Padé approximations to the Grothendieck conjecture on linear differential equations*, Number theory (New York, 1983–84), Lecture Notes in Math., vol. 1135, Springer, Berlin, 1985, pp. 52–100. 37

[CC85c] ———, *Padé approximations and Diophantine geometry*, Proc. Nat. Acad. Sci. U.S.A. **82** (1985), no. 8, 2212–2216. 37

[CDT21] Frank Calegari, Vesselin Dimitrov, and Yunqing Tang, *The unbounded denominators conjecture*, 2021, https://arxiv.org/abs/2109.09040. 7, 8, 9, 11, 12, 13, 14, 15, 38, 40, 42, 43, 44, 60, 62, 66, 69, 70, 71, 72, 74, 77, 78, 79, 80, 81, 103, 133, 134, 196, 199

[CDT24] ———, *Arithmetic holonomy bounds and the irrationality of the 2-adic* $\zeta(5)$, 2024. 13, 67, 71, 94, 127, 192, 199, 206

[Che09] Huayi Chen, *Maximal slope of tensor product of Hermitian vector bundles*, J. Algebraic Geom. **18** (2009), no. 3, 575–603. 96

[Chu79] Gregory V. Chudnovsky, *Formules d'Hermite pour les approximants de Padé de logarithmes et de fonctions binômes, et mesures d'irrationalité*, C. R. Acad. Sci. Paris Sér. A-B **288** (1979), no. 21, A965–A967. 35, 51, 56, 174, 175

[Chu80] ———, *Rational and Padé approximations to solutions of linear differential equations and the monodromy theory*, Complex analysis, microlocal calculus and relativistic quantum theory (Proc. Internat. Colloq., Centre Phys., Les Houches, 1979), Lecture Notes in Phys., vol. 126, Springer, Berlin-New York, 1980, pp. 136–169. 38, 52

[Chu83a] ———, *Number theoretic applications of polynomials with rational coefficients defined by extremality conditions*, Arithmetic and geometry, Vol. I, Progr. Math., vol. 35, Birkhäuser Boston, Boston, MA, 1983, pp. 61–105. 51, 198, 199

[Chu83b] ———, *On the method of Thue–Siegel*, Ann. of Math. (2) **117** (1983), no. 2, 325–382. 35, 50, 51, 56, 175

[Chu05] Wenchang Chu, *Harmonic number identities and Hermite–Padé approximations to the logarithm function*, J. Approx. Theory **137** (2005), no. 1, 42–56. 57

[Coh78] Henri Cohen, *Démonstration de l'irrationalité de $\zeta(3)$ (d'après R. Apéry)*, Séminaire de théorie des nombres de Grenoble **6** (1977–78), VI.1–9. 3, 18

[Coo12] Shaun Cooper, *Sporadic sequences, modular forms and new series for $1/\pi$*, Ramanujan J. **29** (2012), no. 1-3, 163–183. 197

[Dèb85] Pierre Dèbes, *Quelques remarques sur un article de Bombieri concernant le théorème de décomposition de Weil*, Amer. J. Math. **107** (1985), no. 1, 39–44. 190

[Dèb86] ———, *G-fonctions et théorème d'irreductibilité de Hilbert*, Acta Arith. **47** (1986), no. 4, 371–402. 37, 190

[Den50] Jacques Deny, *Les potentiels d'énergie finie*, Acta Math. **82** (1950), 107–183. 125

[DF87] Persi Diaconis and David Freedman, *A dozen de Finetti-style results in search of a theory*, Ann. Inst. H. Poincaré Probab. Statist. **23** (1987), no. 2, 397–423. 59

[DGS94] Bernard Dwork, Giovanni Gerotto, and Francis J. Sullivan, *An introduction to G-functions*, Annals of Mathematics Studies, vol. 133, Princeton University Press, Princeton, NJ, 1994. 8, 18, 52, 53, 57, 189

[DHKK22] Sinnou David, Noriko Hirata-Kohno, and Makoto Kawashima, *Linear forms in polylogarithms*, Ann. Sc. Norm. Super. Pisa Cl. Sci. **XXIII** (2022), no. 5, 1447–1490. 192

[Dir1837] Peter Gustav Lejeune Dirichlet, *Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält*, 1837. 3

[dlVP49] Charles-Jean de la Vallée-Poussin, *Le potentiel logarithmique — balayage et représentation conforme*, 1949. 126

[DT97] Michael Drmota and Robert F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997. 60

[DV01] Lucia Di Vizio, *Sur la théorie géométrique des G-fonctions. Le théorème de Chudnovsky à plusieurs variables*, Math. Ann. **319** (2001), no. 1, 181–213. 52, 189, 192

[Dwo99] Bernard M. Dwork, *On the size of differential modules*, Duke Math. J. **96** (1999), no. 2, 225–239.

[Dys47] Freeman John Dyson, *The approximation to algebraic numbers by rationals*, Acta Math. **79** (1947), 225–240. 192

[DZ14] Simon Dauguet and Wadim Zudilin, *On simultaneous diophantine approximations to $\zeta(2)$ and $\zeta(3)$*, J. Number Theory **145** (2014), 362–387. 85

[Ell06] Richard S. Ellis, *Entropy, large deviations, and statistical mechanics*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1985 original. 61, 94

[Eul1735] Leonhard Euler, *De summis serierum reciprocarum*, 1735. 3

[Fis04] Stéphane Fischler, *Irrationalité de valeurs de zêta (d'après Apéry, Rivoal, ...)*, Astérisque (2004), no. 294, vii, 27–62. 51

[FPLL11a] Ulises Fidalgo Prieto and Guillermo Tomás López-Lagomasino, *Nikishin systems are perfect*, Constructive Approximation **34** (2011), no. 3, 297–356. 51

[FPLL11b] ———, *Nikishin systems are perfect. The case of unbounded and touching supports*, J. Approx. Theory **163** (2011), no. 6, 779–811. 51

[FR03] Stéphane Fischler and Tanguy Rivoal, *Approximants de Padé et séries hypergéométriques équilibrées*, J. Math. Pures Appl. (9) **82** (2003), no. 10, 1369–1394. 51, 192

[FR17] ———, *On the denominators of the Taylor coefficients of G-functions*, Kyushu J. Math. **71** (2017), no. 2, 287–298. 12, 189

[FR18] ———, *Rational approximation to values of G-functions, and their expansions in integer bases*, Manuscripta Math. **155** (2018), no. 3-4, 579–595. 190

[FR21] ———, *Linear independence of values of G-functions, II: outside the disk of convergence*, Ann. Math. Qué. **45** (2021), no. 1, 53–93. 195

[Fug60] Bent Fuglede, *The logarithmic potential in higher dimensions*, Mat.-Fys. Medd. Danske Vid. Selsk. **33** (1960), no. 1, 14. 125, 126

[FW94] Gerd Faltings and Gisbert Wüstholz, *Diophantine approximations on projective spaces*, Invent. Math. **116** (1994), no. 1-3, 109–138. 60

[FW08] Peter J. Forrester and S. Ole Warnaar, *The importance of the Selberg integral*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 4, 489–534. 198, 199

[Gal74] Alexander Ivanovich Galočkin, *Lower bounds of polynomials in the values of a certain class of analytic functions*, Mat. Sb. (N.S.) **95(137)** (1974), 396–417, 471. 189, 191

[Gal75] ———, *Lower bounds of linear forms of the values of certain G-functions*, Mat. Zametki **18** (1975), no. 4, 541–552. 191

[Gal96] ———, *Lower bounds for linear forms of values of G-functions*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1996), no. 3, 23–29, 91, in Russian. 191

[Gol69] Gennady Mikhailovich Goluzin, *Geometric Theory of Functions of a Complex Variable*, Translations of Mathematical Monographs, Vol. 26, American Mathematical Society, Providence, R.I., 1969. 210

[Gro07] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, English ed., Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2007, Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. 59

[GS92] Henri Gillet and Christophe Soulé, *An arithmetic Riemann-Roch theorem*, Invent. Math. **110** (1992), no. 3, 473–543. 98

[Had1899] Jacques Hadamard, *Théorème sur les séries entières*, Acta Math. **22** (1899), no. 1, 55–63. 176, 178

[Hat93] Masayoshi Hata, *Rational approximations to the dilogarithm*, Trans. Amer. Math. Soc. **336** (1993), no. 1, 363–387. 25

[Hat98] ———, *The irrationality of* $\log(1 + 1/q) \log(1 - 1/q)$, Trans. Amer. Math. Soc. **350** (1998), no. 6, 2311–2327. 191, 193

[Her1874] Charles Hermite, *Sur la fonction exponentielle*, 1874. 4, 35, 50, 55

[Her1893] ———, *Sur la généralisation des fractions continues algébriques*, Annali di Matematica, 2$^{\text{e}}$ série, **XXI** (1893), 289–308, Extrait d'une lettre à M. Pincherle. 35, 50

[Her1917] ———, *Oeuvres*, 1917.

[Hil62] Einar Hille, *Analytic function theory. Vol. II*, Introductions to Higher Mathematics, Ginn and Company, Boston, Mass.-New York-Toronto, 1962. 125

[Hoe63] Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30. 61

[HPHP11] Khodabakhsh Hessami Pilehrood and Tatiana Hessami Pilehrood, *Bivariate identities for values of the Hurwitz zeta function and supercongruences*, Electron. J. Combin. **18** (2011), no. 2, Paper 35, 30. 3

[Jac1859] Carl Gustav Jacob Jacobi, *Untersuchungen über die Differentialgleichung der hypergeometrischen Reihe*, J. Reine Angew. Math. **56** (1859), 149–165. 55, 57

[Jag64] Hendrik Jager, *A multidimensional generalization of the Padé table. I–VI*, Indag. Math. **26** (1964), 193–249, Nederl. Akad. Wetensch. Proc. Ser. A **67**. 51, 57

[Kat72] Nicholas M. Katz, *Algebraic solutions of differential equations (p-curvature and the Hodge filtration)*, Invent. Math. **18** (1972), 1–118. 12

[Kol59] Ellis R. Kolchin, *Rational approximation to solutions of algebraic differential equations*, Proc. Amer. Math. Soc. **10** (1959), 238–244. 38, 51

[KZ01] Maxim Kontsevich and Don Zagier, *Periods*, Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 771–808. 3, 161

[Lan66] Serge Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966. 51

[Lan72] Naum S. Landkof, *Foundations of modern potential theory*, Die Grundlehren der mathematischen Wissenschaften, vol. Band 180, Springer-Verlag, New York-Heidelberg, 1972, Translated from the Russian by A. P. Doohovskoy. 125

[Led01] Michel Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, vol. 89, American Mathematical Society, Providence, RI, 2001. 59, 61

[Leg1794] Adrien-Marie Legendre, *Eléments de Géométrie*, 1794. 4

[Lin1882] Ferdinand von Lindemann, *Über die Zahl* π, Math. Ann. **20** (1882), no. 2, 213–225. 3, 4

[Lys18] Vladimir Genrikhovich Lysov, *On Diophantine approximants for the product of logarithms*, 2018, https://keldysh.ru/papers/2018/prep2018_158.pdf (in Russian), pp. 1–20. 191

[Mah53] Kurt Mahler, *On the approximation of* π, Indag. Math. **15** (1953), 30–42, Nederl. Akad. Wetensch. Proc. Ser. A **56**. 51

[Mah68] ———, *Perfect systems*, Compositio Math. **19** (1968), 95–166. 50

[Mah76] ———, *Lectures on transcendental numbers*, Lecture Notes in Mathematics, vol. Vol. 546, Springer-Verlag, Berlin-New York, 1976. 51

[Mah19a] ———, *Ein Beweis der Transzendenz der* P*-adischen Exponentialfunktion*, Doc. Math. (2019), 325–331, Reprint of the original 1933 paper. 21

[Mah19b] ———, *On the approximation of logarithms of algebraic numbers*, Doc. Math. (2019), 527–555, Reprint of the original 1953 paper. 51

[Mai1906] Édmond Maillet, *Introduction à la théorie des nombres transcendants et des propriétés arithmétiques des fonctions*, 1906. 51

[Mai27] Wilhelm Maier, *Potenzreihen irrationalen Grenzwertes*, J. Reine Angew. Math. **156** (1927), 93–148. 25

[Mas16] David W. Masser, *Auxiliary polynomials in number theory*, Cambridge Tracts in Mathematics, vol. 207, Cambridge University Press, Cambridge, 2016. 57, 182

[Mat97] Lutz Mattner, *Strict definiteness of integrals via complete monotonicity of derivatives*, Trans. Amer. Math. Soc. **349** (1997), no. 8, 3321–3342. 126

[Mil82] John Milnor, *Hyperbolic geometry: the first 150 years*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 1, 9–24. 4

[Mil83] ———, *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, Enseign. Math. (2) **29** (1983), no. 3-4, 281–322. 4

[Mil92] Vitali Milman, *Dvoretzky's theorem—thirty years later*, Geom. Funct. Anal. **2** (1992), no. 4, 455–479. 59

[MS86] Vitali Milman and Gideon Schechtman, *Asymptotic theory of finite-dimensional normed spaces*, Lecture Notes in Mathematics, vol. 1200, Springer-Verlag, Berlin, 1986, With an appendix by M. Gromov. 59

[Nag97] Makoto Nagata, *Regular singularities in G-function theory*, Analytic number theory (Kyoto, 1996), London Math. Soc. Lecture Note Ser., vol. 247, Cambridge Univ. Press, Cambridge, 1997, pp. 321–336. 192

[Nai82] Mohan Nair, *A new method in elementary prime number theory*, J. London Math. Soc. (2) **25** (1982), no. 3, 385–391. 199

[Nes88] Yuri V. Nesterenko, *Estimates for the number of zeros of certain functions*, New advances in transcendence theory (Durham, 1986), Cambridge Univ. Press, Cambridge, 1988, pp. 263–269. 53

[Nes96] ———, *Modular functions and transcendence questions*, Mat. Sb. **187** (1996), no. 9, 65–96. 53

[Nes08] ———, *Algebraic independence of* p*-adic numbers*, Izv. Ross. Akad. Nauk Ser. Mat. **72** (2008), no. 3, 159–174. 21

[Nes16] ———, *On Catalan's constant*, Trudy Mat. Inst. Steklova **292** (2016), no. Algebra, Geometriya i Teoriya Chisel, 159–176. 154, 163

[Nes19] ———, *Mahler and transcendence: Effective constructions in transcendental number theory*, Doc. Math. (2019), 123–148. 21

[Nev70] Rolf Nevanlinna, *Analytic Functions*, Die Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York-Berlin, 1970. 91, 210

[Nie1909] Niels Nielsen, *Der eulersche dilogarithmus und seine verallgemeinerungen: Eine monographie*, Nova Acta Leop. **90** (1909), 121–212. 155

[Nik80] Evgeny Mikhailovich Nikišin, *Simultaneous Padé approximants*, Mat. Sb. (N.S.) **113(155)** (1980), no. 4(12), 499–519, 637. 51

[NP01] Yuri V. Nesterenko and Patrice Philippon (eds.), *Introduction to algebraic independence theory*, Lecture Notes in Mathematics, vol. 1752, Springer-Verlag, Berlin, 2001, With contributions from F. Amoroso, D. Bertrand, W. D. Brownawell, G. Diaz, M. Laurent, Yuri V. Nesterenko, K. Nishioka, Patrice Philippon, G. Rémond, D. Roy and M. Waldschmidt. 53

[NS91] Evgeny Mikhailovich Nikišin and Vladimir Nikolaevich Sorokin, *Rational approximations and orthogonality*, Translations of Mathematical Monographs, vol. 92, American Mathematical Society, Providence, RI, 1991, Translated from the Russian by Ralph P. Boas. 51, 125

[Osg85] Charles F. Osgood, *Sometimes effective Thue–Siegel-Roth–Schmidt–Nevanlinna bounds, or better*, J. Number Theory **21** (1985), no. 3, 347–389. 38, 53

[Phi91] Patrice Philippon, *Sur des hauteurs alternatives. I*, Math. Ann. **289** (1991), no. 2, 255–283. 4

[Pól1923] George Pólya, *Sur les séries entières a coefficients entièrs*, Proc. London Math. Soc. (2) **21** (1923), 22–38. 7, 68

[Pól28] ———, *Über gewisse notwendige Determinantenkriterien für die Fortsetzbarkeit einer Potenzreihe*, Math. Ann. **99** (1928), no. 1, 687–706. 8

[Pom69] Christian Pommerenke, *Hankel determinants and meromorphic functions*, Mathematika **16** (1969), 158–166. 8

[PR21] Marc Prévost and Tanguy Rivoal, *Diagonal convergence of the remainder Padé approximants for the Hurwitz zeta function*, J. Number Theory **222** (2021), 346–361. 192

[Pré96] Marc Prévost, *A new proof of the irrationality of $\zeta(2)$ and $\zeta(3)$ using Padé approximants*, J. Comput. Appl. Math. **67** (1996), no. 2, 219–235. 192

[PS98] A. N. Parshin and I. R. Shafarevich, *Number theory. IV*, Encyclopaedia of Mathematical Sciences, vol. 44, Springer-Verlag, Berlin, 1998, Transcendental Numbers, *by Naum I. Fel'dman and Yuri V. Nesterenko*, A translation of *Number theory. 4 (Russian)*, Ross. Akad. Nauk, Vseross. Inst. Nauchn. i Tekhn. Inform., Moscow, Translation by N. Koblitz. 57, 190

[PZ84] Alberto Perelli and Umberto Zannier, *On recurrent mod p sequences*, J. Reine Angew. Math. **348** (1984), 135–146. 13, 17, 37, 44, 206, 209

[Rie38] Marcel Riesz, *Intégrales de Riemann–Liouville et potentiels*, Acta sci. math. Szeged **9** (1938), 1–42. 125, 126

[Riv00] Tanguy Rivoal, *La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs*, C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), no. 4, 267–270. 5, 51

[Riv06] ———, *Nombres d'Euler, approximants de Padé et constante de Catalan*, Ramanujan J. **11** (2006), no. 2, 199–214. 163

[Riv19] ———, *Les E-fonctions et G-fonctions de Siegel*, 2019, p. 77. 190

[Rob68] Raphael M. Robinson, *An extension of Pólya's theorem on power series with integer coefficients*, Trans. Amer. Math. Soc. **130** (1968), 532–543. 68, 196

[Rot55] Klaus F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), 1–20; corrigendum, 168. 39

[Roy90] Ranjan Roy, *The discovery of the series formula for π by Leibniz, Gregory and Nilakantha*, Math. Mag. **63** (1990), no. 5, 291–306. 3

[RR91] Svetlozar Rachev and Ludgar Rüschendorf, *Approximate independence of distributions on spheres and their stability properties*, Ann. Probab. **19** (1991), no. 3, 1311–1337. 59

[RT86] Georges Rhin and Philippe Toffin, *Approximants de Padé simultanés de logarithmes*, J. Number Theory **24** (1986), no. 3, 284–297. 192

[RV96] Georges Rhin and Carlo Viola, *On a permutation group related to $\zeta(2)$*, Acta Arith. **77** (1996), no. 1, 23–56. 85

[RV05] ———, *The permutation group method for the dilogarithm*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **4** (2005), no. 3, 389–437. 25

[RV19] ———, *Linear independence of 1,* Li_1 *and* Li_2, Mosc. J. Comb. Number Theory **8** (2019), no. 1, 81–96. 25

[Sal07] Vladislav Kh. Salikhov, *On the irrationality measure of* $\ln 3$, Dokl. Akad. Nauk **417** (2007), no. 6, 753–755. 29

[SB85] Jan Stienstra and Frits Beukers, *On the Picard–Fuchs equation and the formal Brauer group of certain elliptic* $K3$*-surfaces*, Math. Ann. **271** (1985), no. 2, 269–304. 159

[Sch36] Theodor Schneider, *Über die Approximation algebraischer Zahlen*, J. Reine Angew. Math. **175** (1936), 182–192. 39

[Sch66] Laurent Schwartz, *Théorie des distributions*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, vol. IX-X, Hermann, Paris, 1966, Nouvelle édition, entiérement corrigée, refondue et augmentée. 125

[Shi59] Andrei Borisovich Shidlovskii, *A criterion for algebraic independence of the values of a class of entire functions*, Izv. Akad. Nauk SSSR Ser. Mat. **23** (1959), 35–66. 38, 46, 51

[Shi89] ———, *Transcendental numbers*, De Gruyter Studies in Mathematics, vol. 12, Walter de Gruyter & Co., Berlin, 1989, Translated from the Russian by Neal Koblitz, With a foreword by W. Dale Brownawell. 38, 51

[Sie1921] Carl Ludwig Siegel, *Approximation algebraischer Zahlen*, Math. Z. **10** (1921), no. 3-4, 173–213. 39

[Sie37] ———, *Die Gleichung* $ax^n - by^n = c$, Math. Ann. **114** (1937), no. 1, 57–68. 56, 57

[Sie49] ———, *Transcendental Numbers*, Annals of Mathematics Studies, vol. No. 16, Princeton University Press, Princeton, NJ, 1949. 52

[Sor96] Vladimir Nikolaevich Sorokin, *On the measure of transcendency of the number* π^2, Mat. Sb. **187** (1996), no. 12, 87–120. 51

[Sor16] ———, *On Salikhov's integral*, Trans. Moscow Math. Soc. (2016), 107–126. 29, 85

[Sta23] *In Pascal's triangle without the* 1*s, what is the sum of squares of reciprocals?*, Mathematics Stack Exchange, 2023, https://math.stackexchange.com/q/4828596 (version: 2023-12-17). 3

[Sto74] Kenneth Stolarsky, *Algebraic numbers and Diophantine approximation*, Pure and Applied Mathematics, vol. No. 26, Marcel Dekker, Inc., New York, 1974. 60

[SZ90] Gideon Schechtman and Joel Zinn, *On the volume of the intersection of two* L_p^n *balls*, Proc. Amer. Math. Soc. **110** (1990), no. 1, 217–224. 59

[Thu77] Axel Thue, *Selected mathematical papers of Axel Thue*, Universitetsforlaget, Oslo–Bergen–Tromsø, 1977, with an introduction by Carl Ludwig Siegel. 35, 56, 57

[Thu82] William P. Thurston, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 3, 357–381. 4

[Thu97] ———, *Three-dimensional geometry and topology. Vol. 1*, Princeton Mathematical Series, vol. 35, Princeton University Press, Princeton, NJ, 1997, Edited by Silvio Levy. 4

[vdP80] Alfred van der Poorten, *Some wonderful formulae ... footnotes to Apéry's proof of the irrationality of* $\zeta(3)$, Séminaire Delange-Pisot-Poitou, 20e année: 1978/1979. Théorie des nombres, Fasc. 2 (French), Secrétariat Math., Paris, 1980, pp. Exp. No. 29, 7. 174

[vdP79] ———, *A proof that Euler missed...Apéry's proof of the irrationality of* $\zeta(3)$, Math. Intelligencer **1** (1978/79), no. 4, 195–203, An informal report. 3, 5, 18, 174

[Vio04] Carlo Viola, *The arithmetic of Euler's integrals*, Riv. Mat. Univ. Parma (7) **3*** (2004), 119–149. 85

[Wan04] Julie Tzu-Yueh Wang, *An effective Schmidt's subspace theorem over function fields*, Math. Z. **246** (2004), no. 4, 811–844. 38, 54

[Wei1885] Karl Weierstrass, *Zu Lindemann's Abhandlung: 'Über die Ludolph'sche Zahl".*, Sitzungsberichte der Königlich Preussischen Akademie der Wissen-schaften zu Berlin **5** (1885), no. 2, 1067–1085. 3, 4

[Wei77] André Weil, *Remarks on Hecke's lemma and its use*, Algebraic number theory (Kyoto Internat. Sympos., Res. Inst. Math. Sci., Univ. Kyoto, Kyoto, 1976), Japan Soc. Promotion Sci., Tokyo, 1977, pp. 267–274. 161

[Wir71] Eduard Wirsing, *On approximations of algebraic numbers by algebraic numbers of bounded degree*, 1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969), Proc. Sympos. Pure Math., vol. Vol. XX, Amer. Math. Soc., Providence, RI, 1971, pp. 213–247. 39, 60, 61

[Zag09] Don Zagier, *Integral solutions of Apéry-like recurrence equations*, Groups and symmetries, CRM Proc. Lecture Notes, vol. 47, Amer. Math. Soc., Providence, RI, 2009, pp. 349–366. 1, 6, 10, 18, 159, 161, 162, 163, 164, 199

[Zan09] Umberto Zannier, *Lecture notes on Diophantine analysis*, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], vol. 8, Edizioni della Normale, Pisa, 2009, With an appendix by Francesco Amoroso. 57

[Zan14] Umberto Zannier (ed.), *On some applications of Diophantine approximations*, Quaderni/Monographs, vol. 2, Edizioni della Normale, Pisa, 2014, A translation of Carl Ludwig Siegel's "Über einige Anwendungen diophantischer Approximationen" by Clemens Fuchs, With a commentary and the article "Integral points on curves: Siegel's theorem after Siegel's proof" by Fuchs and Umberto Zannier. 8, 25, 52, 55, 182, 189

[Zha95] Shouwu Zhang, *Positive line bundles on arithmetic varieties*, J. Amer. Math. Soc. **8** (1995), no. 1, 187–221. 98, 131

[Zud01] Wadim Zudilin, *One of the numbers* $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ *is irrational*, Uspekhi Mat. Nauk **56** (2001), no. 4(340), 149–150. 5

[Zud03] ———, *An Apéry-like difference equation for Catalan's constant*, Electron. J. Combin. **10** (2003), Research Paper 14, 10. 163

[Zud14] ———, *Two hypergeometric tales and a new irrationality measure of* $\zeta(2)$, Ann. Math. Qué. **38** (2014), no. 1, 101–117. 85, 192

[Zud17] ———, *A determinantal approach to irrationality*, Constr. Approx. **45** (2017), no. 2, 301–310. 8, 23, 28, 163

Email address: `fcale@math.uchicago.edu`

The University of Chicago, 5734 S University Ave, Chicago, IL 60637, USA

Email address: `dimitrov@caltech.edu`

Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA

Email address: `yungqing.tang@berkeley.edu`

Department of Mathematics, University of California, Berkeley, Evans Hall, Berkeley, CA 94720, USA