

Arithmetic holonomy bounds and effective Diophantine approximation^{*}

Frank Calegari[†], Vesselin Dimitrov[‡] and Yunqing Tang[§]

Abstract. In this paper, we explore several threads arising from our recent joint work on arithmetic holonomy bounds, which were originally devised to prove new irrationality results based on the method of Apéry limits. We propose a new method to address effective Diophantine approximation on the projective line and the multiplicative group. This method, and all our other results in the paper, emerged from quantifying our holonomy bounds in a way that directly yields effective measures of irrationality and linear independence. Applying these to a dihedral algebraic construction, we derive good effective irrationality measures for high order roots of an algebraic number, in an approach that might be considered a multivalent continuation of the classical hypergeometric method of Thue, Siegel, and Baker. A well-known Dirichlet approximation argument of Bombieri allows one to derive from this the classical effective Diophantine theorems, hitherto only approachable by Baker’s linear forms in logarithms or by Bombieri’s equivariant Thue–Siegel method. These include the algorithmic resolution of the two-variable S -unit equation, the Thue–Mahler equation, and the hyperelliptic and superelliptic equations, as well as the Baker–Feldman effective power sharpening of Liouville’s theorem. We also give some other applications, including irrationality measures for the classical $L(2, \chi_{-3})$ and the 2-adic $\zeta(5)$, and a new proof of the transcendence of π . Due to space limitations, a full development of these ideas will be deferred to future work.

1 Effective Diophantine approximation on \mathbf{G}_m . Denote by $h : \mathbf{G}_m(\overline{\mathbf{Q}}) \rightarrow [0, \infty)$ the canonical logarithmic absolute Weil height on the multiplicative group of the field of algebraic numbers, and by $H := \exp(h)$ its exponentiated (“multiplicative”) version. We refer to Bombieri and Gubler’s book [23, § 1.5.7] for the definitions and basic properties of these notions, and in particular [23, § 1.4.3] for the normalization $|\cdot|_v$ we will be using of the absolute value in the place $v \in M_K$ of a number field K . These are the normalizations (see also subsection 1.1) such that, with $\log^+(t) := \max(0, \log |t|)$ and $\alpha \in \overline{\mathbf{Q}}$ is arbitrary, we have $h(\alpha) = \sum_{v \in M_K} \log^+ |\alpha|_v$ independently of a choice of number field K containing α .

A landmark theorem in Diophantine analysis takes on the following minimalistic but structurally robust form.

THEOREM 1.1. *Consider the following data:*

- a number field K ;
- a place v of K ;
- a finitely generated subgroup $\Gamma < K^\times$ of the multiplicative group $\mathbf{G}_m(K) = K^\times$;
- a positive number $\varepsilon > 0$.

From these data, one can extract an effectively computable function $C(K, v, \Gamma, \varepsilon) \in \mathbf{R}$, for which the following Diophantine boundedness property takes place: For every $A \in K^\times$, all solutions $\gamma \in \Gamma$ of the Diophantine inequality

$$(1.1) \quad |1 - A\gamma|_v \leq H(\gamma)^{-\varepsilon}$$

satisfy the effective height bound

$$(1.2) \quad h(\gamma) \leq C(K, v, \Gamma, \varepsilon) (1 + h(A)).$$

^{*}2020 Mathematics Subject Classification. 11J68, 11J72

[†]The University of Chicago, 5734 S University Ave, Chicago, IL 60637, USA. Supported in part by NSF Grants DMS-2001097 and DMS-2450123. (fcale@math.uchicago.edu)

[‡]Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA. (dimitrov@caltech.edu)

[§]Department of Mathematics, University of California, Berkeley, Evans Hall, Berkeley, CA 94720, USA. Supported in part by NSF Grant DMS-2231958 and a Sloan Research Fellowship. (yunqing.tang@berkeley.edu).

One goal of this paper is to outline a new proof of Theorem 1.1. We begin, however, with an abridged history. Theorem 1.1 was first proved by Alan Baker [7] in the Archimedean case (enhancing his original method for lower bounds on linear forms in logarithms in many variables through ideas of Stark and Feldman), and Kunrui Yu [65] in the nonarchimedean case. The effectivity clause is the important one—it is the feature allowing for an algorithmic output of all solutions to many Diophantine equations and inequalities (in the spirit of Hilbert’s Tenth Problem). In contrast, the mere existence of some $C(K, v, \Gamma, \varepsilon) \in \mathbf{R}$ follows easily from Roth’s (or indeed from Siegel’s) ineffective Diophantine exponent of an arbitrary algebraic number, as Gelfond observed based on Siegel’s argument with collecting coset representatives under a high degree isogeny $[r] : \mathbf{G}_m \rightarrow \mathbf{G}_m$. (See, for example, [62, Theorem 1.9] for how the $A = 1$ case follows ineffectively from Roth’s theorem; the general argument for $A \in K^\times$ is similar and implicit in this reduction scheme.)

By a standard and easy argument (see [23, § 5.4.3]), the special case $A = 1$ of Theorem 1.1 (formulated as Theorem 5.4.1 in [23]) provides an effective height bound on the general S -unit equation $x + y = 1$ in two variables, to be solved in the group of S -units $x, y \in \mathcal{O}_{K,S}^\times$ in a number field K , with $S \subset M_K$ signifying an arbitrary finite set of places that contains all Archimedean places, namely:

$$(1.3) \quad h(x) + h(y) \leq \log 16 + 2 \max_{v \in S} \left\{ C \left(K, v, \mathcal{O}_{K,S}^\times, \frac{1}{2|S|} \right) \right\}.$$

In turn, by the classical arguments of Siegel and Mahler (see [23, § 5.3] or [42, § 9.6]), an effective height bound (1.3) leads in theory to an algorithmic resolution of the general Thue–Mahler $F(x, y) \in \mathcal{O}_{K,S}^\times$ and superelliptic $y^m = f(x)$ equations in S -integers $x, y \in \mathcal{O}_{K,S}$, and more broadly to compute the set of integral points on various other (although far from all) affine algebraic curves. Here, $F(x, y) \in K[x, y]$ is an arbitrary homogeneous form and $f(x) \in K[x]$ is an arbitrary polynomial. Furthermore, the linear dependence in Equation (1.2) on the parameter $h(A)$, which in the Archimedean case is formulated and proved as Baker’s Theorem 1.10 in [30, §§ 1, 4, 11, E], also allows one (via the Thue equation, see [30, §§ 4.3, 4.4]) to derive a form of the Baker–Feldman theorem [43] on the effective improvement $\mu_{\text{eff}}(\alpha) \leq \deg \alpha - \epsilon(\mathbf{Q}(\alpha))$ of Liouville’s Diophantine exponent $d = \deg \alpha$ for an arbitrary algebraic number of a degree $d \geq 3$.

The best (known) dependence

$$(1.4) \quad C(K, v, \Gamma, \varepsilon) \ll_{[K:\mathbf{Q}], v, \varepsilon, t} \prod_{i=1}^t (1 + h(\xi_i))$$

on the heights of a set of generators ξ_1, \dots, ξ_t of Γ (apart from an absolute numerical coefficient involved, which however is frequently quite important in practical applications) comes from Baker’s theory of linear forms in logarithms in many variables, cf. [9] or [61, 62] in the Archimedean case, and [66, 67, 68] in the nonarchimedean case. Note that the ultimate Baker–Wüstholz dependency in the form (1.4) eliminates the special role of the coset $A\Gamma$, and $A = 1$ is no longer any loss of generality in this formulation. The form (1.4) has no reason to have any finality; the most optimistic folklore conjecture closely related to *abc* speculates that the product in (1.4) should get replaced by a sum. For simplicity of the exposition, we shall hence stick to the minimalistic essential form of Theorem 1.1, including the distinguished coset parameter A in the shape of Baker’s *Sharpening II* [7], and we abstain from pursuing here the refinement to the form (1.4). A number of different and easier proofs [16, 29], [30, § 4.1], [62, Corollary 10.18] of Theorem 1.1 were obtained using Bombieri’s geometry of numbers argument. The most rudimentary form of Bombieri’s idea [30, Lemma 4.2] goes back to an observation of Stark [58, page 262] which we will use ourselves in section 4. This argument only requires relatively simpler-to-prove bounds [51, 52, 31] for two logarithms, and in particular it does not require the optimal (logarithmic) dependence on the variables of the linear form (or the exponents, as one frequently says). What they *do* require, however, is that they come with Waldschmidt’s optimal range of uniformity in the arguments (inputs) of the logarithms. (This feature was first introduced in [61, *parameter M on page 179*], see also [62, *Theorem 9.1 (ii) with the modified parameter B*].)

Once the reduction to softer bounds on linear forms in two logarithms (given Waldschmidt’s uniformity feature) had been made, it opened up a second and more algebraic path to Theorem 1.1 independent from the analytic theory of linear forms in logarithms. This was carried out by Bombieri in the 1990s, grounded in a Galois equivariant form of the Thue–Siegel principle found by Bombieri, van der Poorten, and Vaaler [18, 24]. A general and completely explicit bound (1.2), which for large enough $h(A)$ is also of the Baker–Wüstholz form (1.4) (although with an inferior dependence on the implicit parameters $[K : \mathbf{Q}]$, v , ε , and $t = \text{rank}(\Gamma)$), can be found in

Bombieri [17] in the Archimedean case, and Bombieri–Cohen [21] and [22] in the nonarchimedean case. Bombieri’s original work [17] was founded upon Viola’s geometric form [60] of Dyson’s nonvanishing lemma, whereupon the introduction of an extrapolation parameter thereafter in [21, 22] denoted “ M ” led in [21] (worked out in detail only in the nonarchimedean case) to a logarithmically improved implicit coefficient in the final estimate of the form (1.4). Finally in [22] the appeal to Viola’s theorem was bypassed and replaced by an elementary Wronskian argument, and the final implicit coefficient in the form (1.4) (once again for $h(A) \gg_{[K:\mathbf{Q}],v,\varepsilon,t} 1$, and worked out there in the nonarchimedean case) was improved some further.

Ultimately, both the Baker–Waldschmidt and the Thue–Siegel (Bombieri) paths to Theorem 1.1 (the only known proofs) remain fairly complicated, and there is no easy path known to such a theorem, even in the special case $A = 1$. This is despite the central role of Theorem 1.1 in modern Diophantine analysis, and after several decades of intensive research (see [30, 8, 10, 54] for a panorama of this subject, and to [19, 20, 39] for an introduction to the paradigm’s history and some speculation about its future in the light of the famous and elusive *abc* conjecture). In this paper, we give a new—and perhaps simpler—proof of Theorem 1.1 by applying our quantitative arithmetic holonomy bounds that we develop below. Explicit bounds will be given in a future paper as we continue to explore the potential of the method we outline here. Two longstanding unsolved problems in the area are to give an effective height bound on the multivariable S -unit equation, and to effectivize Thue’s theorem for general algebraic numbers besides those of the binomial form $\sqrt[r]{a/b}$.

1.1 Notation. As we already said in the introduction, all our conventions and notations for heights and absolute values are the ones found in Bombieri and Gubler’s book [23]. We spell out, in particular, that $|\cdot|_v$ is normalized so that $|p|_v = p^{-\frac{[K_v:\mathbf{Q}_p]}{[K:\mathbf{Q}]}}$ if $v \mid p$, and $|x|_v = |x|_\infty^{\frac{[K_v:\mathbf{R}]}{[K:\mathbf{Q}]}}$ for $v \mid \infty$ and $|\cdot|_\infty$ the Euclidean norm on $K_v \subset \mathbf{C}$. We let $\mathbf{D}_v(R)$ and $\overline{\mathbf{D}}_v(R)$ denote the open and closed unit discs of radius R in \mathbf{C}_v , and $\mathbf{T}_v(R) = \partial\overline{\mathbf{D}}_v(R) := \{|x|_v = R\}$; we sometimes omit the subscript when $v = \infty$, and we also abridge $\mathbf{D}_v := \mathbf{D}_v(1)$, $\overline{\mathbf{D}}_v := \overline{\mathbf{D}}_v(1)$, $\mathbf{T}_v := \mathbf{T}_v(1)$. More generally, we follow the notation of [34]. We will use c_1, c_2, \dots to denote a sequence of easily computable absolute constants, and $c_1(\dots), c_2(\dots), \dots$ will be taken to mean readily computable constructive functions of the data indicated within the brackets. Similarly, the Vinogradov symbols \ll, \asymp , and \gg used without subscripts will imply absolute effective constants, and in general implicit constants that are readily and effectively computable in terms of the data listed in the subscript. The semantical meaning of every occurrence of those symbols is that effective implicit conditions can be written down for which the given statement holds.

2 Arithmetic holonomy bounds: a quantitative theory.

2.1 Holonomy bounds and irrationality. Let us recall the framework in [34, § 1.2, § 2] for proving the irrationality of certain periods η . We start (typically) with $A(x) \in \mathbf{Z}[[x]]$ and $B(x) \in \mathbf{Q}[[x]]$ such that $B(x) - \eta A(x)$ has larger convergence radius than both $A(x)$ and $B(x)$ as power series in $\mathbf{R}[[x]]$. By analogy with [59, 15], we call η an Apéry limit. We pick a holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \rightarrow (\mathbf{C}, 0)$ such that $\varphi^*(B(x) - \eta A(x))$ is meromorphic on $\mathbf{D} \subset \overline{\mathbf{D}}$. We then bound the largest number of $\mathbf{Q}(x)$ -linearly independent power series $f_1, \dots, f_m \in \mathbf{Q}[[x]]$ such that the denominators of their coefficients have the same shape (or better) as those of $B(x)$, and additionally that $\varphi^* f_i$ is meromorphic on \mathbf{D} for all i . If we assume for contradiction that $\eta \in \mathbf{Q}$, then we can take some of these f_i ’s to be $B(x) - \eta A(x)$ and its derivatives, but there may exist (unconditionally) additional such functions. Our holonomy bound (see Theorem 2.1 below) provides an upper bound of m in terms of φ and denominator types when $|\varphi'(0)|$ is sufficiently large comparing to the denominators. We deduce that $\eta \notin \mathbf{Q}$ precisely when we find more functions f_i (including $B(x) - \eta A(x)$ and its derivatives) than the upper bound provided by the following theorem proved in [34, Theorem 2.5.1].

THEOREM 2.1. *Consider two positive integers $m, r \in \mathbf{N}_{>0}$ and an $m \times r$ rectangular array of nonnegative real numbers $\mathbf{b} := (b_{i,j})_{1 \leq i \leq m, 1 \leq j \leq r}$, all of whose columns have the form $0 = b_{1,j} = \dots = b_{u_j,j} < b_{u_j+1,j} = \dots = b_{m,j} =: b_j$, $\forall j = 1, \dots, r$, for some $u_j \in \{0, 1, \dots, m\}$ depending on the column. Let $\sigma_i := b_{i,1} + \dots + b_{i,r}$, $i = 1, \dots, m$ be the i -th row sum, and define $\tau(\mathbf{b}) := \frac{1}{m^2} \sum_{i=1}^m (2i-1)\sigma_i = \sigma_m - \frac{1}{m^2} \sum_{j=1}^r u_j^2 b_j \in [0, \sigma_m]$.*

Consider a holomorphic mapping $\varphi : (\overline{\mathbf{D}}, 0) \rightarrow (\mathbf{C}, 0)$ as above with derivative (conformal size) satisfying the condition $\log |\varphi'(0)| > \tau(\mathbf{b})$. 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

$$(2.1) \quad f_i(x) = \sum_{n=0}^{\infty} a_{i,n} \frac{x^n}{[1, \dots, b_{i,1} \cdot n] \cdots [1, \dots, b_{i,r} \cdot n]}, \quad a_{i,n} \in \mathbf{Z},$$

such that $f_i(\varphi(z)) \in \mathbf{C}[[z]]$ is the germ of a meromorphic function on $|z| < 1$ for all $i = 1, \dots, m$; assume that all f_i are holonomic. Then

$$(2.2) \quad m \leq \frac{\iint_{\mathbf{T}^2} \log |\varphi(z) - \varphi(w)| \mu_{\text{Haar}}(z) \mu_{\text{Haar}}(w)}{\log |\varphi'(0)| - \tau(\mathbf{b})}.$$

Let $\vec{b} := (b_1, \dots, b_r)$ and $\mathcal{V}(\varphi, \vec{b})$ denote the $\mathbf{Q}(x)$ -span of power series f where $\varphi^* f$ is meromorphic on \mathbf{D} and

$$(2.3) \quad f(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{[1, \dots, b_1 \cdot n] \cdots [1, \dots, b_r \cdot n]}, \quad a_n \in \mathbf{Z}.$$

When $\log |\varphi'(0)| > \sum_{j=1}^r b_j$, Theorem 2.1 provides an upper bound of $\dim_{\mathbf{Q}(x)} \mathcal{V}(\varphi, \vec{b})$.

Remark 2.2. In [34, §§ 6–8], we allow more general denominators of f_i , and we have various improvements and alternative bounds. For the rest of this section, we will discuss upper bounds on the irrationality measure of η and use the bound (2.2) as a showcase. We hope it will be clear to the reader how the proof idea of Theorem 2.6 can be used to refine all our qualitative holonomy bounds in [34, §§ 6–8] to include an irrationality measure. By the same token, simpler proofs of the bounds needed for the application to Theorem 1.1 can be more directly obtained with the Perelli–Zannier dynamic box principle technique that we explain in [34, Appendix B].

2.2 Irrationality measures. We recall the definition of irrationality measure.

DEFINITION 2.3. Let v be a place of \mathbf{Q} . The irrationality exponent/measure of $\eta \in \mathbf{Q}_v \setminus \mathbf{Q}$, denoted by $\mu(\eta)$, is the supremum of the set of numbers κ such that $0 < \left| \eta - \frac{p}{q} \right|_v < \frac{1}{q^\kappa}$ is satisfied by infinitely many $(p, q) \in \mathbf{Z} \times \mathbf{Z}_{>0}$.

In other words, for $\kappa > \mu(\eta)$, the above inequality only has finitely many solutions (p, q) as above; thus there exists a constant $C = C(\eta, \kappa) \in \mathbf{R}_{>0}$ such that for all $(p, q) \in \mathbf{Z} \times \mathbf{Z}_{>0}$, we have $\left| \eta - \frac{p}{q} \right|_v > \frac{C}{q^\kappa}$. When η is algebraic of degree $d \geq 2$, Liouville's theorem [54, ch. 1 § 1.6] gives $\kappa = d$ with an effective constant C , and in this case there is a distinction to be made between Definition 2.3 (which, by Roth's theorem, reads $\mu(\eta) = 2$ whenever $2 \leq [\mathbf{Q}(\eta) : \mathbf{Q}] < \infty$), and exponents κ for which an explicit $C(\eta, \kappa)$ can be provided. The infimum of the latter κ , though this not a mathematical definition, is commonly denoted $\mu_{\text{eff}}(\eta)$, and it is the notion featured in all our theorems in this paper (see Remark 2.7). We denote by $\mu_{\text{eff}, K, v}(\eta)$ the straightforward generalization over a number field K and place $v \in M_K$ for v -adic K -rational approximations to an $\eta \in K_v \setminus K$.

Theorem 2.1 extends to a quantitative refinement that includes an explicit upper bound on the irrationality measure $\mu(\eta)$ for an Apéry limit $\eta \in \mathbf{R} \setminus \mathbf{Q}$ as considered in subsection 2.1; for concreteness we formulate our next proposition for the case of the Archimedean place of \mathbf{Q} . Let $\mathcal{V}(\varphi, \vec{b}, \eta)$ denote the $\mathbf{R}(x)$ -span of power series $g \in \mathbf{R}[[x]]$ of form $g(x) = B(x) - \eta A(x)$ such that $\varphi^* g$ is meromorphic on \mathbf{D} and $A(x), B(x)$ satisfy (2.3).

PROPOSITION 2.4. Fix $\varphi : (\overline{\mathbf{D}}, 0) \rightarrow (\mathbf{C}, 0)$ and a denominator type \vec{b} with $|\varphi'(0)| > e^{\sum_{j=1}^r b_j}$. Let

$$m = \dim_{\mathbf{R}(x)} \mathcal{V}(\varphi, \vec{b}, \eta) \geq \dim_{\mathbf{Q}(x)} \mathcal{V}(\varphi, \vec{b}) = \gamma \cdot m, \quad \gamma \in \mathbf{Q} \cap (0, 1].$$

Take $\rho \in (0, 1]$ so that there exists a basis $\{g_i = B_i - \eta A_i\}$ of $\mathcal{V}(\varphi, \vec{b}, \eta)$ where $\varphi^* A_i$ and $\varphi^* B_i$ both individually converge on $|z| < \rho$. Consider \mathbf{b} as in Theorem 2.1 such that A_i, B_i satisfy (2.1). Assume that all g_i are holonomic. Then

$$(2.4) \quad m \leq \frac{\iint_{\mathbf{T}^2} \log |\varphi(z) - \varphi(w)| \mu_{\text{Haar}}(z) \mu_{\text{Haar}}(w)}{\log |\varphi'(0)| - \tau(\mathbf{b}) - (1 - \gamma) (2\mu(\eta)^{-1} - (1 - \gamma)\mu(\eta)^{-2}) \log \frac{1}{\rho}},$$

conditionally on the positive denominator.

If we set $\mu(\eta) \rightarrow \infty$ in (2.4), we recover (2.2). In particular, every qualitative irrationality proof by the holonomy bounds method of Theorem 2.1 automatically refines via (2.4) to an explicit measure of irrationality. It must be said, however, that in this form of the result, the extra coefficient 2 of $\mu(\eta)^{-1}$ in (2.4) means that, when the *usual* Apéry limit works, Proposition 2.4 does *not* recover the most basic standard bound on the irrationality measure. It remains an open problem to give an ε -improvement over the usual irrationality measure in every single case.

Remark 2.5. Consider the general denominator type as in [34, Thm 6.0.2 and 7.0.1], i.e., A_i, B_i are of the form

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

where $\mathbf{e} = (e_1, \dots, e_m) \in \mathbf{N}^m$. Define $\tau^\sharp(\mathbf{e})$ as in [34, §6]. Then we have the following bound on $\mu(\eta)$:

$$(2.5) \quad m \leq \frac{\iint_{\mathbf{T}^2} \log |\varphi(z) - \varphi(w)| \mu_{\text{Haar}}(z) \mu_{\text{Haar}}(w)}{\log |\varphi'(0)| - \tau(\mathbf{b}) - \tau^\sharp(\mathbf{e}) - (1 - \gamma) (2\mu(\eta)^{-1} - (1 - \gamma)\mu(\eta)^{-2}) \log \frac{1}{\rho}}.$$

Proposition 2.4 is a special case, with $K = \mathbf{Q}$ and $S = \{v\}$ a singleton, of Theorem 2.6 below. For our more general setup, we consider a number field K together with a finite subset $S \subset M_K$ of its places, whereby K -rational approximations (sometimes restricted to be from a subfield of K) are to be made simultaneously in all places from S . We organize the targets of the approximation into a “soft” collection $\eta = (\eta_u)_{u \in S}$ of local elements $\eta_u \in K_u$, which in theory do not have to be related, but in practice always have some sort of structural connection, although not always by a rule as simple as to take the u -embedding of an underlying element $\eta \in K$. One central example of this, which in a Kummer equivariant form occurs within the proofs of Propositions 3.3 and 3.4, is similar to the equivariant setup [18, page 75] of Bombieri, van der Poorten, and Vaaler: the approximants β are to be drawn from a subfield $F \subset K$, with K/F being Galois and the set $S \subset M_K$ being $\text{Gal}(K/F)$ -stable, and the assignment $u \rightsquigarrow \eta_u$ transforms according to a given projective representation $\pi : \text{Gal}(K/F) \rightarrow \text{PGL}_2(F)$.

Following the framework of André [2, § VIII] but now combined with axiomatics of Apéry limits, we consider a collection $(\varphi_v)_{v \in M_K}$ of analytic functions on an unspecified open neighborhood of $\overline{\mathbf{D}}_v$, with $\varphi_v(z) = z$ for all but finitely many v ; and a collection $(\rho_u)_{u \in S}$ of radii $\rho_u \in (0, 1]$. These data are to be used for imposing and controlling convergence properties of the formal functions in the Apéry limits method. Though one may consider more general linear independence settings, the focus here on single Apéry limits, but used in a power system $\{1, H, \dots, H^k\}$, compels us to consider only the powers $\eta^j := (\eta_u^j)_{u \in S}$ of the original Apéry limit, along with the implied K -rational approximations β^j . We define $\mathcal{V}(\{\varphi_v\}, \{\rho_u\}, \vec{b}, \eta, \nu)$ to be the formal¹ $K(\eta, x)$ -linear span of all formal power series $g \in K(\eta)[[x]]$ of the form $g(x) = \sum_{j=0}^{\nu} \eta^j A_j(x)$ such that:

1. For all $u \in S$, $\varphi_u^* g \in \mathcal{M}(\mathbf{D}_u)$ is meromorphic on \mathbf{D}_u , and every individual $\varphi_u^* A_j$ is convergent on $|z|_u < \rho_u$.
2. For all $v \in M_K \setminus S$, every individual $\varphi_v^* A_j$ is meromorphic on \mathbf{D}_v .
3. There exists a finite subset $S' \subset M_K$ for which $A_0, \dots, A_\nu \in \sum_{n=0}^{\infty} \frac{x^n \mathcal{O}_{K, S'}}{[1, \dots, b_1 \cdot n] \cdots [1, \dots, b_r \cdot n]}.$

THEOREM 2.6. *Given $(\varphi_v)_{v \in M_K}, \eta = (\eta_u)_{u \in S}, (\rho_u)_{u \in S} \in (0, 1]^S$ as above, consider $r, \mu \in \mathbf{N}$ and, for $\nu = 0, 1, \dots, \mu$, a vector $\vec{b}_\nu \in [0, \infty)^r$, and a list of holonomic (over $K(x)$) elements*

$$g_{\nu,1}, \dots, g_{\nu,m_\nu} \in \mathcal{V}(\{\varphi_v\}, \{\rho_u\}, \vec{b}_\nu, \eta, \nu) \setminus \mathcal{V}(\{\varphi_v\}, \{\rho_u\}, \vec{b}_\nu, \eta, \nu - 1)$$

satisfying conditions (1)–(3) above. Assume the totality $\{g_{\nu,i}\}_{0 \leq \nu \leq \mu, 1 \leq i \leq m_\nu}$ of these $m = m_0 + \dots + m_\mu$ holonomic functions to be $K(\eta, x)$ -linearly independent, and order them in some arbitrary way. Consider for the chosen ordering a denominators capping $m \times r$ -matrix \mathbf{b} obeying the constraint spelled out in Theorem 2.1, and having its k -th row vector dominate entrywise the \vec{b} -vector \vec{b}_ν of the k -th function $g_{\nu,i}$.

¹This we take to mean to consider in every realization $u \in S$ the $K_u(x)$ -linear span of the power series $\sum_{j=0}^{\nu} \eta_u^j A_j(x) \in K_u[[x]]$. A $K(\eta, x)$ -linear dependence is defined to be a $K(\eta, x)$ -linear form which evaluates to a $K_u(x)$ -linear dependence simultaneously in all places $u \in S$.

Consider a vector of exponents $\kappa = (\kappa_u)_{u \in S} \in \mathbf{R}_{>0}^S$ subjected to the inequalities:

$$(2.6) \quad m > \frac{\sum_{v \in M_K, v \nmid \infty} \iint_{\mathbf{T}_v^2} \log |\varphi_v(z) - \varphi_v(w)| \mu_{\text{Haar}} + \sum_{v \in M_K, v \nmid \infty} \sup_{z \in \mathbf{T}_v} \log |\varphi_v(z)|_v}{\sum_{v \in M_K} \log |\varphi'(0)|_v - \tau(\mathbf{b}) - \sum_{u \in S} \log \rho_u^{-1} + \frac{(\sum_{u \in S} \kappa_u - \sum_{\nu=0}^{\mu} \nu m_{\nu}/m)^2}{\sum_{u \in S} \frac{\kappa_u^2}{\log \rho_u^{-1}}}},$$

$$\kappa_u \leq \frac{\log \rho_u^{-1} \sum_{v \in S} \frac{\kappa_v^2}{\log \rho_v^{-1}}}{\sum_{v \in S} \kappa_v - \sum_{\nu=0}^{\mu} \nu m_{\nu}/m}, \quad \forall u \in S.$$

Then there exists a constant $C \in \mathbf{R}_{>0}$ such that the following Diophantine property is in place:

$$(2.7) \quad \text{For all } \beta \in K \text{ with height } h(\beta) > C, \text{ there exists a place } v \in S \text{ where } |\eta_v - \beta|_v \geq H(\beta)^{-\kappa_v}.$$

Remark 2.7. By Northcott's theorem [23, Theorem 1.6.8], (2.7) is another way to say that all but finitely many $\beta \in K$ have some $v = v(\beta)$ where $|\eta_v - \beta|_v \geq H(\beta)^{-\kappa_v}$. As in Theorem 1.1, the point of writing it this way is that C in our proof can readily be made explicit in terms of meromorphic presentations $\varphi_u^* g_{\nu,i} = R_{\nu,i}^u / V_{\nu,i}^u$ and $\varphi_v^* A_{\nu,i,j} = S_{\nu,i,j}^v / W_{\nu,i,j}^v$ with $V_{\nu,i}^v(0) = W_{\nu,i,j}^v(0) = 1$ along with the data $(|\eta_u|_u, \kappa_u, \rho_u)_{u \in S}, \mathbf{b}, \sup_{\mathbf{T}_v} \log |R_{\nu,i}^u|, \sup_{\mathbf{T}_v} \log |V_{\nu,i}^u|, \sup_{\mathbf{T}_v(\rho_v)} \log |S_{\nu,i,j}^u|$, and $\sup_{\mathbf{T}_v(\rho_v)} \log |W_{\nu,i,j}^u|$ for all $u \in S$ and $v \in M_K$ (extending $\rho_v := 1$ to $v \in M_K \setminus S$), in a linear dependence on those four suprema uniformly in terms of the Chudnovsky–Osgood functional bad approximability constant for the system $\{A_{\nu,i,j}\}$, as discussed in [34, § 3.2] and bounded explicitly by Chudnovsky's [34, Theorem 3.2.10] in the case of a differentially closed system.

Remark 2.8. Comparing to (2.2), the extra term subtracted off in the denominator of (2.6) is $\sum_{u \in S} \alpha_u - (\sum_{u \in S} \kappa_u - E)^2 / (\sum_{u \in S} \kappa_u^2 / \alpha_u)$, where $\alpha_u := \log \rho_u^{-1}$, $\gamma_{\nu} := m_{\nu}/m$ with $\sum_{\nu=0}^{\mu} \gamma_{\nu} = 1$, $E := \sum_{\nu=0}^{\mu} \nu \gamma_{\nu} \in [0, \mu]$. The positivity of this term for the case $\sum_{u \in S} \kappa_u > E$ follows by the Cauchy–Schwarz inequality. In the proportionality regime $\kappa_u := \alpha_u T$, the second condition in (2.6) is automatically fulfilled, and the subtracted term reduces to $2/(ET) - (E/T)^2 / (\sum_u \alpha_u)$. This with $T \rightarrow \infty$ is how Theorem 2.6 subsumes Theorem 2.1.

2.3 Proof of Theorem 2.6. We use the framework of Bost's slopes method [25, §§4.1, 4.2]. Without loss of generality upon permuting the columns of \mathbf{b} , the indices u_i in the condition spelled out in Theorem 2.1 satisfy $u_1 \leq u_2 \leq \dots \leq u_r$. In our ordering, let us denote by g_i the i -th function $\{g_{\nu,j}\}$, and by $\nu(i)$ the index for which $g_i \in \mathcal{V}(\{\varphi_v\}, \{\rho_u\}, \vec{b}, \eta, \nu(i)) \setminus \mathcal{V}(\{\varphi_v\}, \{\rho_u\}, \vec{b}, \eta, \nu(i) - 1)$.

We present the argument in its qualitative form: assume for contradiction that the constant C does not exist. This means there is some $\varepsilon > 0$ and an infinite sequence of $\beta \in K$ with $H(\beta) \rightarrow \infty$ along which $|\eta_u - \beta|_u < H(\beta)^{-\kappa_u + \varepsilon}$, simultaneously for all $u \in S$. (We shall let $\varepsilon \rightarrow 0$ at the end of the proof.) Consider the following \mathcal{O}_K -module of rank $m(D+1)$:

$$E_D := \bigoplus_{h=0}^r \bigoplus_{i=u_h+1}^{u_{h+1}} \frac{1}{[1, \dots, u_{h+1} b_{h+1} D] \cdots [1, \dots, u_r b_r D]} y_i \cdot \mathcal{O}_K[1/x]_{\leq D},$$

where the orthogonal direct sum is in the category of Hermitian vector bundles over $\text{Spec } \mathcal{O}_K$: for $v \nmid \infty$, the vectors $y_i, y_i x^{-1}, \dots, y_i x^{-D}$ are equipped with the norm $\|y_i x^k\|_v := \max(|\beta|_v, 1)^{\nu(i)} (\max_{\mathbf{T}_v} \log |\varphi_v|_v)^k$. For $v \mid \infty$, the Hermitian metric is defined (as in [34, §§ 7.3.1, 7.3.3]) for $s \in \mathcal{O}_K[1/x]_{\leq D}$ by $\|y_i s\|_v := \max(|\beta|_v, 1)^{\nu(i)} \|s\|_{\text{BC}}$, where $\|\cdot\|_{\text{BC}}$ signifies the Bost–Charles metric [26, §§ 8.2–8.3]. We use \bar{E}_D to denote E_D with this Hermitian metric. Approximately speaking, we will choose an optimizing D in linear proportion with $h(\beta) = \log H(\beta)$. This allows us to focus on $D \rightarrow \infty$ asymptotic estimates as we will let $h(\beta) \rightarrow \infty$ running along our sequence of counterexamples $\beta \in K$. It will be apparent from the local estimates that the argument is actually effective.

By the arithmetic Hilbert–Samuel theorem and a theorem of Bost and Charles [26, Theorem 5.4.1 and Proposition 5.4.2], we have

$$\widehat{\deg} \bar{E}_D = \left(\frac{m}{2} \left(\sum_{v \mid \infty} \iint_{\mathbf{T}_v^2} \log |\varphi_v(z) - \varphi_v(w)| \mu_{\text{Haar}} + \sum_{v \nmid \infty} \sup_{\mathbf{T}_v} \log |\varphi_v|_v \right) + \sum_{h=1}^r u_h^2 b_h \right) D^2 - \left(\sum_{\nu=0}^{\mu} \nu \gamma_{\nu} \right) h(\beta) mD + o(D^2 + Dh(\beta)).$$

Write $g_i(x) = \sum_{j=0}^{\mu} \eta^j A_{i,j}(x)$. We evaluate E_D on $y_i := \sum_{j=0}^{\mu} \beta^j A_{i,j}(x)$, and thus for $h(\beta) \gg 1$ we obtain an injective homomorphism $\psi_D : E_D \hookrightarrow x^{-D}K[[x]]$. We filter $x^{-D}K[[x]]$ as $x^{-D}K[[x]] \supseteq x^{1-D}K[[x]] \supseteq \dots \supseteq x^{n-D}K[[x]] \supseteq \dots$, and we metricize the graded quotients $x^{n-D}K[[x]]/x^{n+1-D}K[[x]]$ by $\|x^{n-D}\|_v = 1$ for all $v \in M_K$. We define $E_D^{(n)} := \psi_D^{-1}(x^{n-D}K[[x]])$. For each $n \in \mathbf{N}$, the evaluation map ψ_D induces an injective homomorphism $\psi_D^{(n)} : E_D^{(n)}/E_D^{(n+1)} \hookrightarrow x^{n-D}K[[x]]/x^{n+1-D}K[[x]]$. Therefore, $\text{rank}(E_D^{(n)}/E_D^{(n+1)}) \in \{0, 1\}$. Let $\mathcal{V}_D := \{n \in \mathbf{N} : \text{rank}(E_D^{(n)}/E_D^{(n+1)}) = 1\}$. We have $\#\mathcal{V}_D = \text{rank } E_D = m(D+1)$. By the Chudnovsky–Osgood functional bad approximability theorem [34, Theorem 3.2.13], we have $\mathcal{V}_D \subset \{0, 1, \dots, (m+\delta)D\}$ once $\delta > 0$ and $D \gg_{\delta} 1$.

Bost’s slopes inequality in our setting gives $\widehat{\deg} \bar{E}_D \leq \sum_{n \in \mathcal{V}_D} \sum_{v \in M_K} h_v(\psi_D^{(n)})$. We provide an upper bound on each local evaluation height $h_v(\psi_D^{(n)})$ as follows. For the ease of notation, we will adopt the convention that $\frac{0^2}{\log 1 - 1} = 0$ and set $\rho_v := 1$, $\kappa_v := 0$, and $\eta_v := \beta$ for $v \notin S$. Write $h_i(x) := f_i(x) - g_i(x) = \sum_{j=0}^{\mu} (\beta^j - \eta^j) A_{i,j}(x)$. Let \mathcal{L} denote the line bundle $\mathcal{O}([0])$ on $\mathbf{P}_{\mathcal{O}_K}^1$.

LEMMA 2.9. *For $v \mid \infty$, we have (here we can replace ε by 0 if $v \notin S$) the local evaluation height upper bound:*

$$h_v(\psi_D^{(n)}) \leq -n \log |\varphi'_v(0)|_v + D \iint_{\mathbf{T}_v^2} \log |\varphi_v(z) - \varphi_v(w)| \mu_{\text{Haar}} + (-\kappa_v - \varepsilon)h(\beta) + n \log(1/\rho_v)^+ + o(n).$$

Proof. For simplicity of notation, we drop v from the subscript and set $|\beta|_+ := \max(|\beta|, 1)$. An easy reduction, see [34, Lemma 7.4.1], lets us to assume that all $\varphi^* g_i \in \mathcal{O}(\bar{\mathbf{D}})$, as well as all $\varphi^* A_{i,j} \in \mathcal{O}(\bar{\mathbf{D}}_{\rho})$, are holomorphic on their respective closed discs, rather than merely meromorphic on the open discs.

Given $\{P_i\}_{i=1}^m \subset K[1/x]_{\leq D}$ with $\|y_i P_i\| \leq 1$, set $s := \sum_{i=1}^m P_i f_i = c_n x^{n-D} + \dots$, which is a section of the line bundle $\mathcal{L}^{\otimes D}$ on the formal neighborhood of 0 in \mathbf{P}_K^1 ; we write $s_1 := \sum_{i=1}^m P_i g_i$ and $s_2 := \sum_{i=1}^m P_i h_i$. Let $(h_i(\varphi(z)))_{\leq n}$ denote the $O(z^{n+1})$ truncation of the z -power series $h_i(\varphi(z))$. Note that $x^D \cdot s$ is a formal function on \mathbf{P}_K^1 and $\varphi^*(x^D \cdot s) = c_n \varphi'(0)^n z^n + \dots$. The function $F(z) := \varphi^*(x^D \cdot s_1) + \sum_{i=1}^m \varphi^*(x^D P_i) \cdot (h_i(\varphi(z)))_{\leq n}$ is holomorphic on $\bar{\mathbf{D}}$ and has the same lowest order term as $\varphi^*(x^D \cdot s)$. Therefore $\log |z^{-n} F(z)|$ is subharmonic, and

$$\begin{aligned} \log |c_n| + n \log |\varphi'(0)| &\leq \int_{\mathbf{T}} \log |F(z)| \mu_{\text{Haar}} \\ &\leq \int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \{|\varphi^*(x^D P_i) \cdot \varphi^* g_i|, |\varphi^*(x^D P_i) \cdot (h_i(\varphi(z)))_{\leq n}|\} \mu_{\text{Haar}} + O(1) \\ &\leq \int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \{|\varphi^*(x^D P_i)| \cdot |\beta|_+^{\nu(i)}\} \mu_{\text{Haar}} + (-\kappa h(\beta) + n \log(1/\rho))^+ + O(1) \\ &= (-\kappa h(\beta) + n \log(1/\rho))^+ - \int_{\mathbf{T}} \log \|\varphi^* x^{-D}\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \mu_{\text{Haar}} + \int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} + O(1) \\ &= (-\kappa h(\beta) + n \log(1/\rho))^+ + D \iint_{\mathbf{T}^2} \log |\varphi_v(z) - \varphi_v(w)| \mu_{\text{Haar}} + \int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} + O(1). \end{aligned}$$

For the third inequality, note that on \mathbf{T} ,

$$|\varphi^* g_i| \cdot |\beta|_+^{-\nu(i)} = O(1), \quad |(h_i(\varphi(z)))_{\leq n}| \cdot |\beta|_+^{-\nu(i)} = O(|\beta - \eta| \max_{0 \leq l \leq n, 1 \leq i \leq m, 0 \leq j \leq \mu} |b_{i,j,l}|) = O(H(\beta)^{-\kappa+\varepsilon} \rho^{-n}),$$

where $b_{i,j,l}$ is the z^l -coefficient of $\varphi^* A_{i,j}$. At the fourth and the fifth lines, by a mild abuse of notation upon embedding by the given $v : K \hookrightarrow \mathbf{C}$, we use $\bar{\mathcal{L}}$ to denote the line bundle $\mathcal{L}_{\mathbf{C}}$ equipped with the Bost–Charles Hermitian metric as in [34, § 7.3.1] (i.e., we ignore all other places, only focus on $v : K \hookrightarrow \mathbf{C}$). The final equality is a computation due to Bost and Charles performed in [34, (7.3.5), (7.4.3)].

It remains to bound $\int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}}$ under the condition $\log \sup_{1 \leq i \leq m} \|y_i P_i\| \leq 0$:

$$\begin{aligned} \int_{\mathbf{T}} \log \sup_{1 \leq i \leq m} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} &\leq \log \int_{\mathbf{T}} \sup_{1 \leq i \leq m} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} \\ &\leq \log \int_{\mathbf{T}} \sum_{i=1}^m \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} = \log \sum_{i=1}^m \int_{\mathbf{T}} \|\varphi^* P_i\|_{\varphi^* \bar{\mathcal{L}}^{\otimes D}} \cdot |\beta|_+^{\nu(i)} \mu_{\text{Haar}} \\ &\leq \log \sup_{1 \leq i \leq m} \|y_i P_i\| + \log m + o(n) \leq o(n). \end{aligned}$$

□

The estimate for $v \nmid \infty$ is similar, with the subharmonicity of $\log |z^{-n} F(z)|$ being replaced by the maximum principle for $z^{-n} F(z)$. Summing the local estimates we derive for the n -th global evaluation height:

$$\begin{aligned} \sum_{v \in M_K^{\text{fin}}} h_v(\psi_D^{(n)}) &= \sum_{h=1}^r b_h \max\{n, u_h D\} - n \sum_{v \in M_K^{\text{fin}}} \log |\varphi'_v(0)|_v + D \sum_{v \in M_K^{\text{fin}}} \sup_{\mathbf{T}_v} \log |\varphi_v|_v \\ &\quad + \sum_{u \in S} (n \log(1/\rho_u) - (\kappa_u - \varepsilon) h(\beta))^+ + o(n). \end{aligned}$$

We plug our estimates of $\widehat{\deg} \bar{E}_D$ and $h_v(\psi_D^{(n)})$ into the slopes inequality $\widehat{\deg} \bar{E}_D \leq \sum_{n \in \mathcal{V}_D} \sum_{v \in M_K} h_v(\psi_D^{(n)})$. The optimal choice of D for a given $h(\beta)$ is found by minimizing

$$(2.8) \quad \left(\sum_{\nu=0}^{\mu} \nu \gamma_{\nu} \right) m h(\beta) / D + \sum_{u \in S} \int_0^m \max\{0, t \log \rho_u^{-1} - (\kappa_u - \varepsilon) h(\beta) / D\} dt.$$

Set $q := h(\beta) / D$, $\chi_u := \frac{(\kappa_u - \varepsilon)q}{\log \rho_u^{-1}}$ for $\rho_u < 1$, and $\chi_u := 0$ otherwise. Later we will pick q so that $\chi_u \leq m$ for all $u \in S$. Under this condition $\chi_u \leq m$, a direct computation transforms (2.8) into

$$\begin{aligned} &\left(\sum_{\nu=0}^{\mu} \nu \gamma_{\nu} \right) m q + \sum_{u \in S} \left(\frac{m^2 - \chi_u^2}{2} \log \rho_u^{-1} - (\kappa_u - \varepsilon)(m - \chi_u)q \right) \\ &= \left(\sum_{u \in S} \frac{(\kappa_u - \varepsilon)^2}{\log \rho_u^{-1}} \right) \frac{q^2}{2} - \left(\sum_{u \in S} (\kappa_u - \varepsilon) - \sum_{\nu=0}^{\mu} \nu \gamma_{\nu} \right) m q + \left(\sum_{u \in S} \log \rho_u^{-1} \right) \frac{m^2}{2}, \end{aligned}$$

This is a quadratic form in q , with minimum taken at $q := \frac{\sum_{u \in S} (\kappa_u - \varepsilon) - \sum_{\nu=0}^{\mu} \nu \gamma_{\nu}}{\sum_{u \in S} \frac{(\kappa_u - \varepsilon)^2}{\log \rho_u^{-1}}} m$. Note that the condition

$\chi_u < m$ does indeed follow from (2.6) for ε sufficiently small.

Now we let $h(\beta) \rightarrow \infty$ and take $D := \lfloor q h(\beta) \rfloor$. The leading (order D^2) terms in the slopes inequality give

$$\begin{aligned} &\frac{m^2}{2} \left(\sum_{v \in M_K} \log |\varphi'_v(0)|_v - \tau(\mathbf{b}) - \sum_{u \in S} \log \rho_u^{-1} + \frac{(\sum_{u \in S} (\kappa_u - \varepsilon) - \sum_{\nu=0}^{\mu} \nu \gamma_{\nu})^2}{\sum_{u \in S} \frac{(\kappa_u - \varepsilon)^2}{\log \rho_u^{-1}}} \right) \\ &\leq \frac{m}{2} \left(\sum_{v \mid \infty} \iint_{\mathbf{T}_v^2} \log |\varphi_v(z) - \varphi_v(w)| \mu_{\text{Haar}} + \sum_{v \nmid \infty} \sup_{z \in \mathbf{T}_v} \log |\varphi_v|_v \right), \end{aligned}$$

which with $\varepsilon \rightarrow 0$ converges to the thesis of the theorem.

3 A dihedral method. We begin by recalling the classical hypergeometric method [54, ch. 1, § 3], by which Thue originally proved his ineffective finiteness theorem on the integer solutions $(X, Y) \in \mathbf{Z}^2$ of the equation $aX^r - bY^r = c$ for $r \geq 3$ and $a, b, c \in \mathbf{Z} \setminus \{0\}$. The hypergeometric equation with three parameters α, β, γ is the second-order linear ODE

$$x(x-1) \frac{d^2 F}{dx^2} + ((\alpha + \beta + 1)x - \gamma) \frac{dF}{dx} + \alpha \beta F = 0,$$

with one solution, regular at the singular point $x = 0$, being given by the power series

$$F(x) = {}_2F_1 \left[\begin{matrix} \alpha & \beta \\ \gamma \end{matrix}; x \right] := \sum_{k=0}^{\infty} \frac{\alpha(\alpha+1) \cdots (\alpha+k-1) \cdot \beta(\beta+1) \cdots (\beta+k-1)}{\gamma(\gamma+1) \cdots (\gamma+k-1)} \frac{x^k}{k!}.$$

A well-known fact going back to Gauss and Jacobi [48, page 160] is that the Hermite–Padé approximants to the hypergeometric functions of the special form ${}_2F_1 \left[\begin{matrix} \alpha & 1 \\ \gamma \end{matrix}; x \right]$ are all explicitly given in terms of hypergeometric polynomials themselves. This applies in particular to the logarithm function $\log(1-x) = -x \cdot {}_2F_1 \left[\begin{matrix} 1 & 1 \\ 2 \end{matrix}; x \right]$ as well as the binomial functions $(1-x)^\nu = {}_2F_1 \left[\begin{matrix} -\nu & 1 \\ 1 \end{matrix}; x \right]$, and we have:

$$(3.1) \quad \begin{aligned} & {}_2F_1 \left[\begin{matrix} -\nu - n & -m \\ -m - n \end{matrix}; x \right] - (1-x)^\nu \cdot {}_2F_1 \left[\begin{matrix} \nu - m & -n \\ -m - n \end{matrix}; x \right] \\ &= (-1)^m \frac{\binom{n+\nu}{m+n+1}}{\binom{m+n}{m}} {}_2F_1 \left[\begin{matrix} -\nu + m + 1 & n + 1 \\ m + n + 2 \end{matrix}; x \right] x^{m+n+1} = (-1)^m \frac{\binom{n+\nu}{m+n+1}}{\binom{m+n}{m}} x^{m+n+1} + O(x^{m+n+2}). \end{aligned}$$

In a famous paper [4] in 1964 preceding his first general lower bounds on linear forms in logarithms of algebraic numbers, Baker followed the method of Thue and Siegel [57] to establish the explicit Diophantine inequality $|\sqrt[3]{2} - p/q| > 10^{-6} q^{-2.955}$ by considering the sequence of rational approximations to $\sqrt[3]{2}$ obtained from the specializations $m = n$, $\nu = 1/3$, and $x := 3/128$ in (3.1). A crucial point of the numerology is that the linear forms thus specialized in 1 and $\sqrt[3]{1 - 3/128} = (5/8)\sqrt[3]{2}$ have exponentially small denominators, for instance of the form $\binom{2n}{n} 2^{7n} 3^{\lfloor n/2 \rfloor}$. More in line with the Apéry limits method and our holonomy bounds, the generating function of those linear forms is a G -function in the sense of Siegel [69, § VII, page 43], obeying a certain third-order linear homogeneous ODE on the domain $\mathbf{P}^1 \setminus \{0, \infty, ((1 \pm \sqrt{1 - 3/128})/2)^{-2}\}$, overconvergent at the first singularity $((1 + \sqrt{1 - 3/128})/2)^{-2} \approx 1.012$, and, as a result, holomorphic all the way up to the next singularity $((1 - \sqrt{1 - 3/128})/2)^{-2} \approx 28784.766$. Chudnovsky’s improvements in [38] came from a close study of which primes dividing $\binom{2n}{n}$ actually appear in the denominator, and therefore to compute the exact asymptotics of the true denominator. In more recent times, Bennett has refined the hypergeometric method, improving [11] Baker’s example to the clean bound $|\sqrt[3]{2} - p/q| > q^{-2.5}/4$, and establishing in collaboration with de Weger the striking theorem [12, 13] that for every $r \geq 3$ and $a, b \in \mathbf{Z}$ the original Thue equation $aX^r - bY^r = 1$ has not more than one solution in positive integers $(X, Y) \in \mathbf{N}^2$ (in particular, by bridging the gap to the wholly complementary range of applicability of the theory of linear forms in logarithms).

3.1 Multivalence. The behavior of the generating functions of the $m = n$ diagonal specializations of the Hermite–Padé forms (3.1)—namely, that they are holonomic and overconvergent at the smallest non-zero singularity—exemplifies the mechanism of overconvergence for Apéry limits described in subsection 2.1. Following [34], we seek to continue the traditional hypergeometric method by considering the analytic continuations of these functions beyond their convergence discs. Ideally, we can find conformally large, possibly multivalent analytic mappings $\varphi : (\overline{\mathbf{D}}, 0) \rightarrow (\mathbf{C}, 0)$, and then apply our holonomy bounds. To obtain better multivalence properties, it turns out to be advantageous to ignore the arithmetic denominator savings in [38], and instead take the generating function of (3.1) after first clearing through the binomial coefficient. For $\nu \in \mathbf{Q}$ —as it turns out—we obtain an algebraic (dihedral) generating function on the domain $y \in \mathbf{P}^1 \setminus \{((1 \pm \sqrt{1-x})/x)^2, \infty\}$ ($y = 0$ is no longer a singularity!). We have the clean hypergeometric polynomial generating function identity

$$\sum_{m,n=0}^{\infty} {}_2F_1 \left[\begin{matrix} -\nu - n & -m \\ -m - n \end{matrix}; x \right] \binom{m+n}{m} y^m z^n = \frac{(1-xy)^\nu}{1-y-z+xyz}.$$

On the $m = n$ diagonal, this gives, by a standard Cauchy integral (over a “vanishing cycle” [44, 41]):

PROPOSITION 3.1. For $\nu \in \mathbf{Q}$ rational of denominator r , the functions

$$(3.2) \quad \begin{aligned} A_\nu(x, y) &:= \sum_{n=0}^{\infty} \sum_{k=0}^n \left\{ \frac{(n-\nu)(n-1-\nu) \cdots (k+1-\nu)}{(n-k)!} \binom{n+k}{k} (-x)^{n-k} \right\} y^n \\ &= \sum_{n=0}^{\infty} {}_2F_1 \left[\begin{matrix} -\nu - n & -n \\ -2n \end{matrix}; x \right] \binom{2n}{n} y^n \in \mathbf{Z}[[y, xy/r^2]], \\ B_\nu(x, y) &:= A_{-\nu}(x, y) = \sum_{n=0}^{\infty} {}_2F_1 \left[\begin{matrix} \nu - n & -n \\ -2n \end{matrix}; x \right] \binom{2n}{n} y^n \in \mathbf{Z}[[y, xy/r^2]] \end{aligned}$$

are algebraic functions of y over $\mathbf{Q}(x)$, with Galois group the spherical triangle group of signature $(2, 2, r)$, that is D_r , if r is odd, and $D_{r/2}$, if r is even. They form a \mathbf{C} -basis of $y = 0$ formal solutions to the second-order linear ODE of “associated Legendre kind” on $y \in \mathbf{P}^1 \setminus \{((1 \pm \sqrt{1-x})/x)^2, \infty\}$:

$$(3.3) \quad (1 - 4y + 2xy + x^2y^2) \frac{d^2 f}{dy^2} + 3(x^2y + x - 2) \frac{df}{dy} + x^2(1 - \nu^2)f = 0.$$

Another \mathbf{C} -basis of solutions to (3.3) are the basic dihedral functions

$$(3.4) \quad \frac{\cos\left(\nu \arccos \frac{x^2y+x-2}{2\sqrt{1-x}}\right)}{\sqrt{1-4y+2xy+x^2y^2}} \quad \text{and} \quad \frac{\sin\left(\nu \arccos \frac{x^2y+x-2}{2\sqrt{1-x}}\right)}{\sqrt{1-4y+2xy+x^2y^2}}.$$

For $|x| \leq 1$, the particular linear combination $B_\nu(x, y) - (1-x)^\nu A_\nu(x, y)$ spans the one-dimensional \mathbf{C} -vector space of solutions of (3.3) that are regular (“overconverge”) at the singularity $y = ((1 - \sqrt{1-x})/x)^2$. The functions $A_\nu(x, y), B_\nu(x, y)$ of (3.2) have the property that, at every finite place $v \in M_{\mathbf{Q}(x, \prod_{p|r-1} p^{1/(p-1)})}^{\text{fin}}$ when $x \in \overline{\mathbf{Q}}$ is algebraic, they map the closed v -adic disc $|y|_v \leq \max(1, |x|_v)^{-1} |r \prod_{p|r} p^{1/(p-1)}|_v$ to the closed unit disc $|Y|_v \leq 1$.

The local monodromies of (3.3) are $\mathbf{Z}/2\mathbf{Z}$ at the finite singularities $y = ((1 \pm \sqrt{1-x})/x)^2$, and $\mathbf{Z}/r\mathbf{Z}$ at $y = \infty$. We take up David and Gregory Chudnovskys’ arithmetic algebraization method [37] for cases of Grothendieck’s p -curvature conjecture (as extended by André [2, §§ 1.2–1.4] and Bost [25] when the monodromy group is *a priori* virtually solvable). The exponential map to a commutative algebraic group gives rise [55] to certain entire meromorphic resolutions φ (generally in several variables) of growth order ≤ 2 . One then exploits these resolutions to prove algebraicity (under the running hypotheses) in a manner similar to that of section 2. In our situation, we have an *a priori* algebraic function $H = B - \eta A$ with holonomic rank 2, but one of a high algebraicity degree $\in \{r, 2r\}$, with an algebraic Apéry limit of the form $\eta = \sqrt[r]{a}$ (if we specialize $x := 1 - a$ with $a \in \overline{\mathbf{Q}}$); and we can take the entire map $\varphi \in \mathcal{O}(\mathbf{C})$ to be holomorphic in just one single variable, and with growth order 1 (for A and B simultaneously), or $1/2$ (for the overconvergent combination $H := B - \eta A$). This pair of maps is universal for the general (“spherical triangle type $(2, 2, \infty)$ ”) ODE on $\mathbf{P}^1 \setminus \{\alpha, \beta, \infty\}$ having order 2 local monodromies around the two finite singularities α and β . The latter map—crucially for our application of growth order strictly < 1 —has never apparently been considered in these circle of problems before. In these arrangements, even though both our functions as well as our Apéry limits are *a priori* algebraic, we continue to get striking conclusions when we apply the Diophantine extensions in section 2 of the holonomy bounds to the case of the system $\{1, H, H^2, \dots, H^k\}$, for an optimal choice of the power $k < r$. We can write down explicit entire holomorphic maps φ that apply to every local system of ramification type $(2, 2, \bullet)$ on $\mathbf{P}^1 \setminus \{\alpha, \beta, \infty\}$, as follows:

LEMMA 3.2. Let $\alpha \neq \beta$, and let $f \in \mathbf{C}[[x]]$ be a holomorphic function germ which analytically continues as a meromorphic function along all paths in $\mathbf{P}^1 \setminus \{\alpha, \beta, \infty\}$, with monodromy of order 2 around both α and β .

1. $\psi_{\alpha, \beta}^* f$ is meromorphic on \mathbf{C} , where $\psi_{\alpha, \beta}(z) = \frac{\alpha + \beta}{2} \cdot \left(1 - \cosh\left(\frac{z}{\sqrt{\alpha\beta}}\right)\right) + \sqrt{\alpha\beta} \cdot \sinh\left(\frac{z}{\sqrt{\alpha\beta}}\right) = z + \dots$
2. Suppose moreover that f extends analytically at $z = \alpha$ along some path γ from 0 to α . Then, for some choice of u with $u := \frac{1}{2} \log \frac{\sqrt{\beta} + \sqrt{\alpha}}{\sqrt{\beta} - \sqrt{\alpha}}$ and $\tanh^2(u) = \alpha/\beta$, the pullback $\varphi_{\alpha, \beta}^* f$ is meromorphic on \mathbf{C} , where

$$(3.5) \quad \varphi_{\alpha,\beta}(z) = \alpha \left(1 - \frac{\sinh \left(\sqrt{u(u - z \tanh(u)/\alpha)}^2 \right)}{\sinh^2(u)} \right) = z + \dots$$

The entire holomorphic maps $\psi_{\alpha,\beta}$ and $\varphi_{\alpha,\beta}$ have growth orders 1 and $1/2$, respectively.

If $|\alpha| < |\beta|$, and the path γ lies within the disc of radius $|\alpha|$, then one can take $u = \tanh^{-1}(\sqrt{\alpha/\beta})$ to be given by the corresponding convergent power series. Lemma 3.2 can be verified by an elementary but unenlightening calculation, so instead we give some hints as to where these formulas come from. For the first claim, we can pull back via a quadratic map ramifying at α and β and reduce to the case $\mathbf{P}^1 \setminus \{1, \infty\}$ with monodromy \mathbf{Z} and universal covering map $\tilde{\psi}(z) = 1 - e^{-z}$. For the second claim, there is a unique quadratic map $h : \mathbf{P}^1 \rightarrow \mathbf{P}^1$ with $h(z) = z + O(z^2)$ that ramifies at $\{\beta, \infty\}$, so $h^{-1}(\infty) = \{\infty\}$ and the order 2 singularity at β is resolved; explicitly, $h(z) = z - z^2/(4\beta)$. Let $h^{-1}(\alpha) = \{\tilde{\alpha}, \tilde{\beta}\}$. The overconvergence along γ pulls back to an overconvergence along a path $\tilde{\gamma}$ from 0 to $\tilde{\alpha}$ (which distinguishes $\tilde{\alpha}$ from $\tilde{\beta}$). Moreover, the monodromy of h^*f is once again of type $(2, 2, \bullet)$. One can now successively iterate these quadratic maps, and the limit converges—in some neighborhood of $z = 0$ —to $\varphi_{\alpha,\beta}$ (where the choice of γ determines u).

3.2 An effective irrationality type for high order roots. Whereas the $x = -1/n$ specialization to $\sqrt[r]{1-x}$ in (3.1) directly constructs a sequence of rational approximations to $\sqrt[r]{1+1/n}$ which converge fast enough to supply an effective irrationality measure $\mu \rightarrow 2$ in the limit $n \gg_r 1$, historically it was not easy to obtain any nontrivial effective irrationality measure for $\sqrt[r]{2}$ for large r . This had to wait until 1971 with the first of Baker’s *Sharpening* series of papers [5]. The asymptotic bound $\mu_{\text{eff}}(\sqrt[r]{2}) \ll \log r$ in [6] is, up to the absolute constant, still the state-of-art today. But Theorem 2.6 with the dihedral algebraic construction work together to directly supply a sub-Liouville effective bound $\mu_{\text{eff}}(\sqrt[r]{2}) \ll \sqrt{r(\log r)^3}$. This irrationality measure is essentially of Siegel’s strength, but it is effective, and applies just as well to any binomial algebraic number $\sqrt[r]{a/b}$. In a way, our argument effectivizes Mahler’s explicit Hermite–Padé treatment [53] of Siegel’s [56] ineffective exponent $2\sqrt{r}$ for the particular case of binomial numbers $\sqrt[r]{a/b}$ with $0 < a < b$. Thue, Siegel, and Mahler worked with the explicit type I Hermite–Padé form directly for the extended binomial system $\{1, (1-x)^{1/r}, \dots, (1-x)^{k/r}\}$ and, with an optimal choice of $k < r$, specialized $x := 1 - aq_0^r/bp_0^r \approx 0$ for a hypothetical exceptionally good “anchoring” approximation $p_0/q_0 \approx \sqrt[r]{a/b}$. By Baker’s theorem, such a “Thue ghost lever” p_0/q_0 is known to *not* exist, and this is the ultimate source of the ineffectivity in the whole Thue paradigm. In contrast, we work with the “dual” system $\{1, H, \dots, H^k\}$ where H is the dihedral algebraic generating function of the Hermite–Padé forms of $\{1, (1-x)^{1/r}\}$, specialized directly at the requisite point $x = 1 - a/b \in (0, 1) \cap \mathbf{Q}$. The growth properties of the entire multivalent maps of Lemma 3.2 (2) are restricted to a suitable choice of radius, and input into Proposition 2.4. The ensuing holonomy bound enables us to make a choice of $k = m - 1$ closely similar to Siegel’s (and, by extension, Mahler’s) optimizing choice, but now the Diophantine bound is effective as we have used the multivalence of $\varphi_{\alpha,\beta}$, in place of the ghost Thue lever p_0/q_0 , for increasing the overconvergence as far as possible.

Concretely in our $\sqrt[r]{2}$ example, consider $H(y) := B_{1/r}(-1, y) - 2^{1/r}A_{1/r}(-1, y) \in \mathbf{Q}[[y]] + 2^{1/r}\mathbf{Q}[[y]]$ from Proposition 3.1, where the corresponding ODE has singularities at $\alpha = 3 - 2\sqrt{2}$ (the overconvergent singularity) and $\beta = 3 + 2\sqrt{2}$. With $K = \mathbf{Q}$, we have $\mathbf{b} = \mathbf{0}$, and the p -adic convergence radii have $R_p = 1$, for $p \nmid r$, and $\prod_{p|r} R_p^{-1} = r \prod_{p|r} p^{1/(p-1)}$. Hence $\sum_{p \in M_{\mathbf{Q}}^{\text{fin}}} \log R_p = -\log r - \sum_{p|r} \frac{\log p}{p-1} \geq -\log r - \log \log r - O(1)$. At the Archimedean place, we select the map $\varphi_{\infty}(z) = \varphi(z) := \varphi_{\alpha,\beta}(Rz)$ where $\varphi_{\alpha,\beta}$ is given by (3.5). Conditionally on $k < r$ and the positive denominator, and making use both of the explicit formula (3.5) and its low asymptotic growth order of $1/2$, our holonomy bound reads

$$(3.6) \quad k + 1 \ll \frac{\sqrt{R}}{\log R - \log(r \log r) - (c_0 k / \mu_{\text{eff}}(\sqrt[r]{2})) \log R}.$$

The choices $R := c_1 r \log r$ and $k := c_2 \sqrt{r \log r}$ give an effective measure of irrationality $\mu_{\text{eff}}(\sqrt[r]{2}) \ll \sqrt{r(\log r)^3}$.

This computation extends directly to the K -relative irrationality measures of arbitrary roots $\sqrt[r]{a}$ from elements $a \in K$, and lower bounds on $|\sqrt[r]{a} - b|$ can then be interpreted as lower bounds for linear forms $\log a - r \log b$ in

two logarithms. But to get to Theorem 1.1 by Bombieri's elementary geometry of numbers reduction in section 4, we need a bound that has Waldschmidt's feature of remaining asymptotically sub-Liouville in the uniformity range $r \gg 1 + h(a)$, as in Bombieri's [17, Theorem 1]. The best such bounds [61, 46, 36] have the form $|\log a - r \log b|_v \gg_v -d^4 h \max(1, \log(r/h))$ with $d = [K : \mathbf{Q}]$ and $h = 1 + h(a)$, but any effective bound of the shape $\gg_{K,v} -r/g(r/h)$ with $g(t) \rightarrow \infty$ is sufficient. (The papers [17, 21, 22] on the Thue–Siegel principle reach $g(t) = \sqrt{\log t}$, $g(t) = t/(\log t)^7$, and $g(t) = t/(\log t)^5$, respectively.) Our method with applying the arithmetic holonomy bounds to Proposition 3.1 leads quite straightforwardly to a bound with $g(t) = t^{\frac{1}{2d}}(\log t)^{-3}$, at least under the assumption — entirely harmless for the application to Theorem 1.1 — that r factorizes suitably.

The above example and computation, contained in the next proposition as the special case $K = \mathbf{Q}$, $l = 1$, and $a = 2$, extends with Waldschmidt's uniformity as follows. In the following, we consider $a \in K$ in our arbitrary number field K , and we write $d := [K : \mathbf{Q}]$.

PROPOSITION 3.3 (Archimedean case). *Let $l, s \in \mathbf{N}$ and $r = ls$. For every place $u \mid v$ of $K(\sqrt[l]{a})$,*

$$(3.7) \quad \log |\sqrt[l]{a} - b|_u \gg_d - (1 + h(a^{1/r}) + h(b)) e^{h(a)/l} l s^{\frac{d}{d+1}} (\log s)^{\frac{2d+1}{d+1}},$$

holds for all $b \in K$, provided $s \geq c_3$. In particular,

$$(3.8) \quad \mu_{\text{eff}, K, u}(\sqrt[l]{a}) \ll_d e^{h(a)/l} l s^{\frac{[K:\mathbf{Q}]}{[K:\mathbf{Q}]+1}} (\log s)^{\frac{2[K:\mathbf{Q}]+1}{[K:\mathbf{Q}]+1}}.$$

Proof. We describe the basic idea, whose proper detail is deferred to a subsequent publication. By a standard reduction, we may replace a by $a^{\pm m}$ for some $1 \leq m \leq 6$ so that $|1 - a^{1/l}|_u \leq 1$. Apply Proposition 3.1 to the exponent $\nu := 1/s$ and the specialization $x := 1 - a^{1/l}$ of height $h(a)/l + O(1)$. With $H := B_{1/s}(1 - a^{1/l}, y) - \sqrt[l]{a} A_{1/s}(1 - a^{1/l}, y) \in K[a^{1/l}][\![y]\!] + \sqrt[l]{a} K[a^{1/l}][\![y]\!]$, we apply Theorem 2.6 to the system $\{1, H, \dots, H^k\}$, changing the variable notation x of section 2 to our present variable y , and working over the number field $K' := K(a^{1/l})$ to replace the field denoted K there. The extension K'/K has degree at most l . At each place $w \in M_{K'}$ extending v , we select $\eta_w := \zeta \sqrt[l]{a}$ for the l -th root of unity $\zeta = \zeta_w$ for which the w -place of $\zeta \sqrt[l]{a}$ is equal to the u -place of $\sqrt[l]{a}$. Here again, $k < s$ is a parameter to optimize. With $\alpha, \beta \in \mathbf{C}$ the dihedral ODE singularities, we take the map $\varphi_w(z) := \varphi_{\alpha, \beta}(R^2 z)$ of Lemma 3.2, for $w \mid v$, and the map $\varphi_w(z) := \psi_{\alpha, \beta}(Rz)$, for the other Archimedean places $w \nmid v$. This time the holonomy quotient (3.6) reads (upon a positive denominator):

$$(3.9) \quad k + 1 \ll \frac{R}{\frac{d+1}{d} \log R - \log(s \log s) - h(a)/l - (c_4 kl / \mu_{\text{eff}, K, u}(\sqrt[l]{a})) \log R}.$$

The similar choices of R and k as in the $\sqrt[3]{2}$ illustrative example now lead directly to the bound (3.8). The uniform dependence on the height $h(a)$ is traced back to the dependence of the constant C in Remark 2.7. \square

In the nonarchimedean case we have a closely similar bound, now based on an Archimedean multivalence in Lemma 3.2 together with the special circumstance of Teichmüller representatives: $\lim_{n \rightarrow \infty} a^{p^n} \in \mu_{p^\infty}(K_v)$, if p is the residual characteristic of K_v .

PROPOSITION 3.4 (Nonarchimedean case). *Let p be the residual characteristic of v and $h \in [1, r/e]$ a parameter. Consider $r = ls$ with $l \equiv 1 \pmod p$, where $s \asymp \sqrt{r/h}$ and $l \asymp \sqrt{rh}$. Then, for all $a \in K$ with $h(a) \leq h$, and at every place $u \mid v$ of $K(\sqrt[l]{a})$,*

$$(3.10) \quad \mu_{\text{eff}, K, u}(\sqrt[l]{a}) \ll_{v, d} h(r/h)^{\frac{2d-1}{2d}} \log^2(r/h).$$

More precisely, for all $b \in K$, the condition $r \geq c_5 h$ implies

$$(3.11) \quad \log |\sqrt[l]{a} - b|_u \gg_{v, d} - (1 + h(\sqrt[l]{a}) + h(b)) h(r/h)^{\frac{2d-1}{2d}} \log^2(r/h).$$

Proof. We indicate only a brief guideline. A standard reduction allows us to assume that $(1 - a)/p$ is an algebraic p -integer. Write $q := p^n$ with n a parameter, and apply Proposition 3.1 to the exponent $\nu := 1/(qs)$ and the specialization $x := 1 - a^{q/l}$, with $|x|_w \leq |q|_w$ for all $w \mid p$, and of height $qh(a)/l + O(1)$. As in subsection 5.2 below, the multivalence input will be Archimedean with a choice of radius R , this time applied to a map

from Lemma 3.2 (1) at all Archimedean places. Now the holonomy quotient reads (upon a positive denominator):

$$(3.12) \quad k + 1 \ll_{v,d} \frac{R + d^{-1} \log q}{\log R + d^{-1} \log q - \log(s \log s) - (\log p)/(p-1) - qh(a)/l - (c_6 kl/\mu_{\text{eff},K,u}(\sqrt[d]{a})) \log q},$$

Select $q \asymp s \asymp \sqrt{r/h}$, so that $qh(a)/l \ll 1$, and $R \asymp (s \log s)/q^{1/d} \asymp s^{1-1/d} \log s$. With $k := \lfloor c_7 s^{1-1/d} \log s \rfloor$ the holonomy bound (3.12) amounts to (3.10). The refinement to (3.11) is once again routine from tracking the constant C in Remark 2.7, as we will do in proper detail in a subsequent paper. \square

3.3 A new proof of the transcendence of π and other logarithms. If we take the $\nu \rightarrow 0$ limit of (3.3), we obtain an ODE whose solutions are given by $A(x, y) := (1 - 4y + 2xy + x^2y^2)^{-1/2}$ and $B(x, y) := 2A(x, y) \log \left(1 - x \left(1 + xy - \sqrt{1 - 4y + 2xy + x^2y^2}\right)/2\right)$, and such that $H := B(x, y) - \log(1 - x)A(x, y)$ is the y -generating series of the diagonal Hermite–Padé approximants of the logarithm function [34, § 3.3.7]. The domain $y \in \mathbf{P}^1 \setminus \{((1 \pm \sqrt{1-x})/x), \infty\}$ and overconvergence property of H are exactly the same as in the discussion with the algebraic dihedral functions in subsection 3.1, but now H is a transcendental function with an infinite dihedral monodromy $D_\infty \cong (\mathbf{Z}/2\mathbf{Z}) * (\mathbf{Z}/2\mathbf{Z})$, and correspondingly $B \in \mathbf{Q}[x][[y]]$ has $[1, \dots, n]$ type denominators. Examining the analog of the holonomy quotients (3.9) and (3.12) leads in this context to a new proof of the transcendence of the complex and p -adic logarithms $\log a$ at all algebraic arguments $a = 1 - x \in \overline{\mathbf{Q}}^\times \setminus \{1\} \hookrightarrow \mathbf{C}$ or $a \in \overline{\mathbf{Q}}^\times \setminus \mu_{p^\infty} \hookrightarrow \mathbf{C}_p$ with $|1 - a|_p < 1$. We illustrate this on the case of π (with $x = 2$). Replacing y by iz , we apply the dihedral method to $H(z) := B(z) - \pi A(z)$ on $z \in \mathbf{P}^1 \setminus \{-1/2, 1/2, \infty\}$, where:

$$(3.13) \quad A(z) := (1 - 4z^2)^{-1/2} \in \mathbf{Z}[[z]], \quad B(z) := -2A(z) \arcsin(2z) \in \sum_{n \geq 0} \frac{z^n \mathbf{Z}}{[1, \dots, n]}.$$

If π were contained by some number field K of some degree d , the system $\{1, H, \dots, H^k\}$ would fit Theorem 2.6 with type $[1, \dots, n][1, \dots, n/2] \cdots [1, \dots, n/k]$, which has $\tau(\mathbf{b}) = \sum_{j=1}^k 1/j = \log k + O(1)$. At this point, taking again the maps $\varphi_{-1/2, 1/2}(R^2 z)$ at one (the “overconvergent”) embedding $K \hookrightarrow \mathbf{C}$ and $\psi_{-1/2, 1/2}(Rz)$ at the $d-1$ other complex embeddings, the—qualitative, for simplicity—holonomy quotient bound similar to (3.9) becomes $k + 1 \ll R / \left(\frac{d+1}{d} \log R - \sum_{j=1}^k 1/j \right)$, provided a positive denominator. With the choice of radius $R := (4k)^{\frac{d}{d+1}}$, this yields a contradiction for $k \gg_d 1$.

The alternate specialization $x = -1$ recovers exactly the approximations used by [1, 59] to produce an explicit irrationality measure for $\log 2$. In contrast, the $x = 2$ specialization considered above does not even directly give a sequence of rational numbers converging to π , but yet still leads to the transcendence of π , and even to explicit (if not impressive) irrationality measures such as $\mu(\pi) \leq 15.086$.

4 Conclusion of the new proof of Theorem 1.1. In Propositions 3.3 and 3.4 we constructed a monotonic function $g = g_{K,v} : \mathbf{R}_{\geq 0} \rightarrow \mathbf{R}_{>1}$ which can be made fully explicit, depending only on K and v and having $\lim_{x \rightarrow \infty} g(x) = \infty$, such that, for all $r \in \mathbf{N}$ possessing a factor $l \asymp 1 + h(a)$, resp. a factor $l \asymp \sqrt{r(1 + h(a))}$ congruent to 1 mod p , the following lower bound in two logarithms holds uniformly for all places $u \mid v$ of $K(\sqrt[d]{a})$ and all elements $a, \eta \in K$:

$$(4.1) \quad \log |1 - a^{1/r} \eta|_u \geq - \left(\log 2 + h(a^{1/r}) + h(\eta) \right) r / g(r/h(a)).$$

Construct $c_8(K, v, \varepsilon) \in \mathbf{R}_{\geq 1}$ so that $g_{K,v}(x) \geq 2/\varepsilon$ for all $x \geq c_8(K, v, \varepsilon)$. Then, for all exponents r that meet our *ad hoc* factorizability condition for (4.1), we derive:

$$(4.2) \quad a, \eta \in K \text{ with } r \geq c_8(K, v, \varepsilon)h(a) \implies |1 - a^{1/r} \eta|_u \geq H(a)^{-\varepsilon/2} (2H(\eta))^{-\varepsilon r/2}.$$

The next type of argument to conclude Theorem 1.1 is due to Bombieri [17, § 8] and is now standard. The simplest form of [17, Lemma 4] is:

LEMMA 4.1. *Consider $n_1, \dots, n_t \in \mathbf{Z}$ and two integer parameters $Q, N \in \mathbf{N}$. There exists a positive integer r of the form $r = Q!N/q$ for some $q \in \{1, \dots, Q\}$ and rational integers $p_1, \dots, p_t \in \mathbf{Z}$ satisfying*

$$(4.3) \quad |n_i - r p_i| \leq r Q^{-1/t}, \quad i = 1, \dots, t.$$

Proof. Dirichlet's approximation theorem applied to the vector $(n_1/Q!N, \dots, n_t/Q!N)$ and with Q as the Dirichlet parameter gives a simultaneous rational approximation $(p_1/q, \dots, p_t/q)$ with some common denominator $q \in \{1, \dots, Q\}$ for which (4.3) holds with $r := Q!N/q \in \mathbf{N}$. \square

Remark 4.2. Of course, one can replace $Q!$ in this argument by the lowest common multiple $[1, \dots, Q]$. Bombieri and Cohen used a slightly more involved version [17, Lemma 4], [21, § 6] with weights, which for large $h(A)$ leads to the bounds with (1.4). We suppress this feature for our introductory purposes in this paper.

Consider an arbitrary $\gamma = \zeta \xi_1^{n_1} \dots \xi_t^{n_t}$ as in Theorem 1.1, where $\zeta \in \Gamma_{\text{tors}}$ is a root of unity and ξ_1, \dots, ξ_t are generators for the free part $\Gamma/\Gamma_{\text{tors}}$. To bound the height of γ means to bound the largest absolute value of an exponent n_i . With the r and the p_1, \dots, p_t coming from Lemma 4.1, there exists a decomposition

$$\gamma = a_0 \eta^r, \quad \eta := \xi_1^{p_1} \dots \xi_t^{p_t}, \quad a_0 := \zeta \xi_1^{n_1 - r p_1} \dots \xi_t^{n_t - r p_t}, \quad \text{with} \quad h(a_0) \leq c_9(\Gamma) r Q^{-1/t}.$$

We arrive at a coset representative $a := Aa_0 \in A\Gamma$ with height satisfying

$$(4.4) \quad |h(a) - h(A)| \leq h(a_0) \leq c_9(\Gamma) r Q^{-1/t}.$$

In Lemma 4.1, we select

$$(4.5) \quad Q = Q(K, v, \Gamma, \varepsilon) := \lceil (2c_8(K, v, \varepsilon))^t c_9(\Gamma)^t \rceil,$$

getting by (4.4) the bound $h(a) \leq h(A) + r/(2c_8(K, v, \varepsilon))$. Upon changing the coset representative A of $A\Gamma$, for the purpose of proving Theorem 1.1 we may and do assume that $h(A) \geq L$, with another and arbitrary constructive function $L := L(K, v, \Gamma, \varepsilon)$ of the data which we are free to select later. Choose $\alpha := \sqrt[t]{a}$ in a place $u \mid v$ of $L := K(\sqrt[t]{a})$ to be the branch closest to the K -rational element $\eta^{-1} \in \Gamma \subset K$. Then $(\alpha\eta)^r = A\gamma$. To supply the requisite exponent factorization $r = ls$, we will take the parameter N in Lemma 4.1 to be a prime power $N = g^m$, where $g \in \{2, 3\}$ is 2 unless v is nonarchimedean of residual characteristic 2, and $g = 3$ in the latter case. Define N to be the smallest power of g that exceeds $2c_8(K, v, \varepsilon)h(A)/(Q-1)!$. Our remark regarding $h(A) > L$ then allows us to assume $N > Q!^2$. This in turn ensures that $r = NQ!/q$ does indeed possess both requisite factorizations. The bound (4.1) is therefore in place, and (4.2) follows.

Our (or rather Bombieri's) choices of Q and N were selected precisely to supply the condition $r \geq c_8(K, v, \varepsilon)h(a)$ for (4.2). First, (4.4) with the choice (4.5) give $h(a) \leq h(A) + r/(2c_8(K, v, \varepsilon))$. Then the requisite bound $r \geq c_8(K, v, \varepsilon)h(a)$ follows from the implication

$$(4.6) \quad r \geq (Q-1)!N \geq 2c_8(K, v, \varepsilon)h(A) \implies r \geq c_8(K, v, \varepsilon)(h(A) + r/(2c_8(K, v, \varepsilon))) \geq c_8(K, v, \varepsilon)h(a),$$

In the direction opposite to (4.2), as in [17, Equation (8.4)], we have

$$(4.7) \quad |1 - \alpha\eta|_u < 2^r |1 - A\gamma|_v \leq 2^r H(\gamma)^{-\varepsilon} \leq 2^{Q!N} H(A)^\varepsilon H(A\gamma)^{-\varepsilon} = 2^{Q!N} H(A)^\varepsilon H(\alpha\eta)^{-\varepsilon r},$$

under our assumption (1.1) in the theorem-under-proof. Together these conditions bound $h(\alpha\eta) \leq c_{10}(K, v, \Gamma, \varepsilon)$, therefore $h(A\gamma) \leq c_{10}(K, v, \Gamma, \varepsilon)r \leq c_{10}(K, v, \Gamma, \varepsilon)Q!N$. Hence ultimately $h(\gamma) \leq h(A) + h(A\gamma) \leq h(A) + c_{10}(K, v, \Gamma, \varepsilon)Q!N \leq c_{11}(K, v, \Gamma, \varepsilon)(1 + h(A))$ due to the choice of N . We arrive at $C(K, v, \Gamma, \varepsilon) := c_{11}(K, v, \Gamma, \varepsilon)$.

5 Irrationality measures: Examples. We finish with two illustrative examples of Theorem 2.6 and its variants. The first is an irrationality measure for $L(2, \chi_{-3})$, making one of the main irrationality results of [34] effective (see Remark 2.7). The second is the irrationality of $\zeta_2(5) \in \mathbf{Q}_2$, which in early 2020 had been our first joint result in this collaboration.

5.1 An irrationality measure for $L(2, \chi_{-3})$. We can apply Proposition 2.4 (in the form Remark 2.5) to give an explicit irrationality measure for $L(2, \chi_{-3})$ using the explicit template $\varphi(z) = h(\psi(z))$ of [34, Def A.4.2]. The overconvergent singularity in this setting corresponds to $-1/72$. The unique preimage x of $-1/72$ of h in the image of $\psi(z)$ is computed in ([34, Lemma A.4.4]); the ρ with $\psi(\rho) = x$ and thus $\varphi(\rho) = -1/72$ is

$$\rho \sim 0.000049668 - 0.000070341i, \quad |\rho^{-1}| < 11614.$$

Precisely 7 of our $m = 14$ functions are genuine (Lemma [34, 12.1.1]), and so $\gamma = 1/2$. Thus (2.5) becomes

$$(5.1) \quad 14 < \frac{11.845}{\log\left(256 \cdot \frac{5448339453535586608000000000}{8658833407565631122430056127}\right) - \left(\frac{27}{80} + \frac{191}{49}\right) - \log(11614) \left(\frac{\kappa - (1/4)}{\kappa^2}\right)}$$

As $\kappa \rightarrow \infty$, the RHS converges to $13.9938 \dots < 14$ (cf. [34, A.5.1]), which shows that $L(2, \chi_{-3}) \notin \mathbf{Q}$. From (5.1), we deduce a contradiction as soon as κ is sufficiently large. Explicitly, we deduce:

THEOREM 5.1. *For all $p/q \in \mathbf{Q}$ except possibly a finite and computable list, we have $\left|L(2, \chi_{-3}) - \frac{p}{q}\right| > \frac{1}{q^{24781}}$.*

The same theorem (and proof) holds with $L(2, \chi_{-3})$ replaced by any \mathbf{Q} -linear combination of $L(2, \chi_{-3})$ and π^2 .

5.2 The irrationality of $\zeta_2(5) \in \mathbf{Q}_2$. Let \mathcal{W}^+ be the space of continuous homomorphisms $\kappa : \mathbf{Z}_p^\times \rightarrow \mathbf{C}_p^\times$ with $\kappa(-1) = 1$. The p -adic zeta function $\zeta_p(\kappa)$ is a rigid analytic function on \mathcal{W}^+ away from the point $\kappa = 1$, at which it has a simple pole. If κ has the form $x \mapsto \chi(x)x^k$ for an integer k , then $\zeta_p(\kappa) = L_p(\chi, 1 - k)$, where $L_p(\chi, k)$ is the Kubota–Leopoldt p -adic L -function [47]. If κ has the form $x \mapsto x^{2k}$ for an integer k , then we write $\zeta_p(1 + 2k) := \zeta_p(\kappa)$. For complex s , write $\zeta^*(s) = (1 - p^{-s})\zeta(s)$ for the p -deprived zeta function, that is, the usual Riemann zeta function with the local factor at p removed. For a negative integer k , there is an equality

$$(5.2) \quad \zeta_p(1 + 2k) = \zeta^*(1 + 2k)$$

which should be interpreted as meaning that both values (*a priori* elements of \mathbf{C}_p and \mathbf{C} respectively) are both inside \mathbf{Q} and are equal. For a *positive* integer k , the continuity of $\zeta_p(\kappa)$ together with (5.2) implies the equality

$$(5.3) \quad \zeta_p(1 + 2k) = \lim_{m \rightarrow k} \zeta^*(1 + 2m),$$

where the limit is over $m \in \mathbf{Z}$ subject to the condition that $2m \equiv 2k \pmod{p-1}$ and the topology on \mathbf{Z} is induced from the inclusion $\mathbf{Z} \subset \mathbf{Z}_p$. One can view (5.3) as interpolating the Kummer congruences on the Bernoulli numbers. The characterization of $\zeta_p(1 + 2k)$ relevant for us is that, for any $2k \in \mathbf{Z} \setminus \{0\}$, it is the unique p -adic number for which the formal power series

$$(5.4) \quad E_{-2k}^* := \frac{\zeta_p(1 + 2k)}{2} + \sum_{n=1}^{\infty} \left(\sum_{d|n, p \nmid d} d^{-2k-1} \right) q^n \in \mathbf{Q}_p + q\mathbf{Q}[[q]]$$

is the q -expansion of an overconvergent p -adic modular form of level $\Gamma_0(p)$ and weight $-2k$ ([40, §2.2]). The specializations E_{2k}^* when k is a *positive* integer are just the classical p -stabilized Eisenstein series of level $\Gamma_0(p)$; they have coefficients in \mathbf{Q} by (5.2). We now prove that the 2-adic zeta value $\zeta_2(5) \in \mathbf{Q}_2$ is irrational.

THEOREM 5.2. *For sufficiently large p or q , we have $\left|\zeta_2(5) - \frac{p}{q}\right|_2 > \frac{1}{\max(|p|, |q|)^{20}}$.*

Proof. We begin with the same setup as [33], where it is proved that $\zeta_2(3) \notin \mathbf{Q}$. The new input is to consider a non-trivial template $\varphi(z)$ at the Archimedean place. We have an identification $X_0(2) \cong \mathbf{P}_x^1$ with the Hauptmodul

$$(5.5) \quad x = x(q) := \frac{\Delta(2\tau)}{\Delta(\tau)} = q \prod_{n=1}^{\infty} (1 + q^n)^{24},$$

where the ordinary locus of $X_0(2)$ is given by the two components $|x|_2 \leq 1$ and $|x|_2 \geq 2^{12}$ and the complementary annulus is the supersingular locus. We may formally invert (5.5) to obtain the expansion

$$q = x - 24x^2 + 852x^3 - 35744x^4 + \dots \in x + x^2\mathbf{Z}[[x]].$$

We now consider (for $p = 2$) the following pair of Eisenstein series (5.4):

$$E_{2k}^* \in \mathbf{Q} + x\mathbf{Z}[[x]], \quad E_{-2k}^* \in \frac{\zeta_2(1 + 2k)}{2} + x\mathbf{Q}[[x]].$$

We have $\tau(E_{2k}^*(x)) = 0$ in the x -coordinate and $\tau(E'_{-2k}(x)) = 2k + 1$ in the x -coordinate, where $E'_{-2k} \in x\mathbf{Q}[[x]]$ is E_{-2k}^* with the constant term omitted. Suppose for contradiction that $\zeta_2(2k+1) \in \mathbf{Q}$, so that $E_{-2k}^* \in \mathbf{Q}[[x]]$. It follows that the overconvergent modular function $H := E_{2k}^* E_{-2k}^*$ for $\Gamma_0(2)$ has rational coefficients and satisfies $\tau(H) = 2k + 1$. One finds that radii $R_p = 1$ for $p \notin \{2, \infty\}$. For $p = 2$, Buzzard's theorem [32, Thm 5.2] gives analyticity of E_{2k}^* and thus of H on the union of the component of the ordinary locus containing ∞ and the entire supersingular locus. This region is precisely the disk $|x|_2 < R_2$ of radius $R_2 = 2^{12}$: hence this is (a lower bound for) the 2-adic convergence radius of the power series $H(x)$. The Archimedean convergence radius of the power series $H(x)$ is $1/64$. For $k = 1$, the usual Apéry argument gives that $\zeta_2(3) \notin \mathbf{Q}$, and this is essentially the content of [33]. For $k = 2$, it is crucial that we exploit the analytic continuation of $H(x)$ to a region beyond this disc. In particular, $H(x) = H(x(q))$ is analytic on the larger region $\{|q| < 1\}$, which has conformal radius 64 times the naïve disc in the coordinate x . As in [34, §A], the goal is to choose a simply connected region $\Omega \subset D(0, 1)$ so that $x(q)$ restricted to Ω is not too large and yet the conformal radius of Ω is not too small. Without trying too hard to optimize Ω , we initially choose Ω to be the circle: $\Omega = \Omega^{\text{circ}} = \left\{ z \in D(0, 1), \left| z + \frac{2}{5} \right| \leq \frac{3}{5} \right\}$, and thus (composing $x(q)$ with the standard uniformization $D(0, 1) \rightarrow \Omega^{\text{circ}}$ sending 0 to 0) the choice

$$\varphi(z) = x\left(\frac{z}{2z+3}\right) = \frac{z}{2z+3} \prod_{n=1}^{\infty} \left(1 + \left(\frac{z}{2z+3}\right)^n\right)^{24}.$$

with $|\varphi'(0)| = 1/3$. By Kontsevich and Zagier ([49], the proof of Fact 1 in section 2.3), the power series $H(x)$ satisfies a linear inhomogeneous differential equation over $\mathbf{Q}(x)$ of the exact minimal order $5 = 2k + 1$, so that the functions $1, H(x), H'(x), H''(x), H'''(x)$ and $H^{(iv)}(x)$ exhibit *six* $\mathbf{Q}(x)$ -linearly independent elements of the space $\mathcal{V}(\varphi_2(z) = 2^{-12}z, \varphi_{\infty}(z) = \varphi(z); (\vec{5}))$. With $R_p = 1$ for $p \neq 2$ and $R_2 = 2^{12}$, and with $r = 1$, $\mathbf{b} = (0, 5, 5, 5, 5, 5)$

and $\tau(\mathbf{b}) = \frac{1}{36} \sum_{i=2}^6 5(2i-1) = \frac{35}{36} \cdot 5$, our holonomy bound Theorem 2.1 implies that

$$(5.6) \quad 6 \leq m \leq \frac{\iint_{\mathbf{T}^2} \log |\varphi(z) - \varphi(w)| + \sum_{\ell} \log R_p}{\log |\varphi'(0)| + \sum_p \log R_p - \tau(\mathbf{b})} = \frac{2.13322\dots + 12 \log 2}{-\log 3 + 12 \log 2 - 175/36} = 4.43206\dots < 6,$$

which is the desired contradiction. We now turn to the irrationality measure. Of the $m = 6$ functions, only one lies in $\mathbf{Q}[[x]]$. Hence $\gamma = 1/6$. With $\rho_2 = 2^{-12}$, we get an irrationality measure for any κ satisfying (cf. (5.6))

$$(5.7) \quad 6 > \frac{2.13322\dots + 12 \log 2}{-\log 3 + 12 \log 2 - 175/36 - 12 \log(2)^{\frac{2(1-\gamma)\kappa - (1-\gamma)^2}{\kappa^2}}},$$

or $\kappa \sim 22.0724$. This can be improved with a more involved choice of Ω . For example, consider the alternate choice $\varphi(z) = x(\psi(z))$ with the lune contour [34, (A.1.1)] $\psi : D(0, 1) \xrightarrow{\sim} \Omega^{\text{lune}} \subset \overline{D(0, 2/3)}$ given explicitly by:

$$\psi(z) = \frac{2}{3}h\left(-z, \frac{2}{5}\right) = \frac{2}{3} \cdot \frac{(5^2 + 2^2)}{2(5^2 - 2^2)} \left(1 + z - \sqrt{1 - \frac{2(5^4 - 6 \cdot 5^2 2^2 + 2^4)}{(5^2 + 2^2)^2} z + z^2}\right) = \frac{14}{29} \cdot z + \dots$$

This leads to the estimate $\kappa < 19.7439 < 20$, where the analogue of (5.7) is

$$6 < \frac{3.92881\dots + 12 \log 2}{(\log 14 - \log 29) + 12 \log 2 - 175/36 - 12 \log(2)^{\frac{2(1-\gamma)\kappa - (1-\gamma)^2}{\kappa^2}}}.$$

We bound the value of the Bost–Charles integral above by the rearrangement integral [34, Prop 8.1.13] which is easier to compute rigorously. \square

Remark 5.3. The proof of Theorem 5.2 neither requires the full force of the holonomy bounds in [34] nor even of their precursors in [35]. Indeed, the proof goes through (with the same φ_{∞} corresponding to Ω^{circ}) where the Bost–Charles integral term in the numerator is replaced by $\max \log |\varphi(z)|$ and $\tau(\mathbf{b})$ is replaced by $\tau = 5$. The corresponding computation for (5.6) is

$$(5.8) \quad 6 \leq m \leq \frac{\max \log |\varphi(z)| + \sum_{\ell} \log R_p}{\log |\varphi'(0)| + \sum_p \log R_p - \tau} = \frac{\log((1/5) \prod_{n=1}^{\infty} (1 + (1/5)^n)^{24}) + 12 \log 2}{-\log 3 + 12 \log 2 - 5} = 5.52667\dots < 6.$$

We proved such bounds (together with the application to $\zeta_2(5)$) in 2020 (see [34, §15.7]), but we put the argument aside in order to work towards the main theorems of [34]; we apologize for the delay! Quite recently, Lai, Sprang, and Zudilin [50] have found a completely new proof of Theorem 5.2 more in the spirit of Apéry’s original argument by working with a differential equation with better convergence properties. The corresponding local system in their case appears to have a genuinely different geometric origin, coming from a moduli space of Calabi–Yau 3-folds rather than a symmetric power of the Picard–Fuchs equation of a modular curve. Neither of these results, however, have any direct bearing on the (ir)rationality of the classical $\zeta(5)$.²

Acknowledgments. As discussed in [34], our work on arithmetic holonomy bounds owes an obvious debt to the mathematics of the Chudnovskys [37, 38], André [2], Bost [25], and Bost–Charles [26].

References

- [1] K. ALLADI AND M. L. ROBINSON, *Legendre polynomials and irrationality*, J. Reine Angew. Math., 318 (1980), pp. 137–155.
- [2] Y. ANDRÉ, *G-Functions and Geometry*, no. E13 in Aspects of Mathematics, Friedr. Vieweg Sohn, Braunschweig, 1989.
- [3] Y. ANDRÉ, *Galois theory, motives and transcendental numbers*, in Renormalization and Galois theories, vol. 15 of IRMA Lect. Math. Theor. Phys., Eur. Math. Soc., Zürich, 2009, pp. 165–177.
- [4] A. BAKER, *Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers*, Quart. J. Math. Oxford Ser. (2), 15 (1964), pp. 375–383.

²One may wonder why one can prove the irrationality of $\zeta_2(5) \in \mathbf{Q}_2$ but not the irrationality of $\zeta(5) \in \mathbf{R}$. We give a heuristic explanation why the first problem is easier. Let $\mathcal{MT}(\mathbf{Z})$ denote the Tannakian category of mixed Tate motives over \mathbf{Z} . Associated to $\mathcal{MT}(\mathbf{Z})$ is a \mathbf{Q} -algebra of real (fixed by complex conjugation) motivic periods \mathcal{P} [27]. The comparison theorem between Betti cohomology and de Rham cohomology (over \mathbf{C}) induces a period map $\mathcal{P} \rightarrow \mathbf{R}$, and Grothendieck’s period conjecture implies that this map is injective [3]. A consequence of this conjecture is that the \mathbf{Q} -algebra generated by $\zeta(2)$ and the odd values $\zeta(2k+1)$ for $k \geq 1$ is free on these generators; equivalently, that these values are all transcendental and moreover algebraically independent. What is perhaps surprising at first is that the p -adic analogue of this story behaves somewhat differently [14, 45]. One of the most basic manifestations of this difference is that the p -adic comparison theorem naturally places the p -adic analogue of $2\pi i$ as an element in B_{dR}^+ which maps to zero in \mathbf{C}_p . In particular, the only sensible definition of $\zeta_p(2)$ returns $\zeta_p(2) = 0$ ([63, Rem 3.7]). One nonetheless conjectures [64] that the \mathbf{Q} -algebra $\mathbf{Q}(\zeta_p(3), \zeta_p(5), \zeta_p(7), \dots)$ is free (on the obvious generators). Let us consider how the vanishing of $\zeta_p(2)$ affects approaches to irrationality. The first step in any Apéry type scheme is to produce a motivic local system where the corresponding L -value arises as a period. One issue, when trying to construct such local systems is that various periods get “mixed” together and it is hard to separate the period one is particularly interested in. (One interesting discussion of these issues from a geometric point of view can be found in [28].) The point is now that, p -adically, the vanishing of $\zeta_p(2)$ makes it easier to realize the odd zeta values on their own. Relevantly for us, the construction in [33] directly generalizes to produce an ODE over $\mathbf{P}^1 \setminus \{0, -1/2^6, \infty\}$ associated to $\zeta_2(2k+1)$ for any k ; we know of no simple analogous construction for $\zeta(2k+1) \in \mathbf{R}$. As another example; the ODE for the 2-adic Catalan’s constant in [33] has the form $\mathcal{L}F = 0$ over $\mathbf{P}^1 \setminus \{0, -1/2^4, \infty\}$ where

$$\mathcal{L} = x(1+16x)^2 \frac{d^2}{dx^2} + (1+16x)^2 \frac{d}{dx} - 4.$$

Up to scaling this is a hypergeometric equation which one can solve explicitly, namely, with:

$$(5.9) \quad A(x) = (1+16x)^{1/2} \cdot {}_2F_1 \left[\begin{matrix} 1/2 & 1/2 \\ 1 \end{matrix}; -16x \right], \quad C(x) = (1+16x)^{1/2} \cdot {}_2F_1 \left[\begin{matrix} 1/2 & 1/2 \\ 1 \end{matrix}; 1+16x \right].$$

The holomorphic solution $B(x) \in \mathbf{Q}[[x]]$ to the non-homogenous equation $\mathcal{L}F = 1+16x$ with $B(0) = 0$ and $B'(0) = 1$ is of denominator type $\tau = [1, 2, \dots, n]^2$. Another solution to this equation $\mathcal{L}F = 1+16x$ is given explicitly by

$$D(x) = -\frac{1+16x}{4} \cdot {}_3F_2 \left[\begin{matrix} 1 & 1 & 1 \\ 3/2 & 3/2 \end{matrix}; 1+16x \right].$$

Over the real numbers, $D(x)$ is singular at $x = 0$, and we have the equality

$$(5.10) \quad D(x) = B(x) - \frac{G}{2} \cdot A(x) - \frac{\pi^2}{16} \cdot C(x),$$

where $G \in \mathbf{R}$ is now the usual Catalan’s constant. In contrast, 2-adically, we see that $D(x)$ converges for $|1+16x|_2 < 2^4$, or for $|x|_2 < 2^8$, and so it is certainly regular at $x = 0$. Now the 2-adic avatar of (5.10) simplifies to the formula

$$(5.11) \quad D(x) = B(x) - \frac{G_2}{2} \cdot A(x),$$

in which $G_2 \in \mathbf{Q}_2$ is the 2-adic Catalan constant. Thus (5.11) realizes G_2 as an Apéry limit. The key point is now that (5.10) is precisely the “same” equation as (5.11) as long as we interpret the 2-adic meaning of $(2\pi i)^2$ to be zero, but Apéry’s method only allows one to deduce that $G_2 \in \mathbf{Q}_2 \setminus \mathbf{Q}$, but not that $G \in \mathbf{R} \setminus \mathbf{Q}$. That said, while this is an argument to explain why the p -adic zeta values of small weight should be easier to study than their Archimedean counterparts, there is also the contrasting remark that we can’t as yet even rule out the possibility that $\zeta_p(3) = 0$ for all sufficiently large p , let alone prove that all those values are irrational.

- [5] A. BAKER, *A sharpening of the bounds for linear forms in logarithms*, Acta Arith., 21 (1972), pp. 117–129.
- [6] A. BAKER, *Recent advances in transcendence theory*, Trudy Mat. Inst. Steklov., 132 (1973), pp. 67–69, 264.
- [7] A. BAKER, *A sharpening of the bounds for linear forms in logarithms. II*, Acta Arith., 24 (1973), pp. 33–36.
- [8] A. BAKER, *Transcendental number theory*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2022. With an introduction by David Masser, Reprint of the 1975 original.
- [9] A. BAKER AND G. WÜSTHOLZ, *Logarithmic forms and group varieties*, J. Reine Angew. Math., 442 (1993), pp. 19–62.
- [10] A. BAKER AND G. WÜSTHOLZ, *Logarithmic forms and Diophantine geometry*, vol. 9 of New Mathematical Monographs, Cambridge University Press, Cambridge, 2007.
- [11] M. A. BENNETT, *Effective measures of irrationality for certain algebraic numbers*, J. Austral. Math. Soc. Ser. A, 62 (1997), pp. 329–344.
- [12] M. A. BENNETT, *Rational approximation to algebraic numbers of small height: the Diophantine equation $|ax^n - by^n| = 1$* , J. Reine Angew. Math., 535 (2001), pp. 1–49.
- [13] M. A. BENNETT AND B. M. M. DE WEGER, *On the Diophantine equation $|ax^n - by^n| = 1$* , Math. Comp., 67 (1998), pp. 413–438.
- [14] A. BESSER, P. BUCKINGHAM, R. DE JEU, AND X.-F. ROBLLOT, *On the p -adic Beilinson conjecture for number fields*, Pure Appl. Math. Q., 5 (2009), pp. 375–434.
- [15] F. BEUKERS, *Irrationality proofs using modular forms*, 1987. Journées arithmétiques (Besançon, 1985).
- [16] Y. BILU AND Y. BUGEAUD, *Démonstration du théorème de Baker–Feldman via les formes linéaires en deux logarithmes*, J. Théor. Nombres Bordeaux, 12 (2000), pp. 13–23.
- [17] E. BOMBIERI, *Effective Diophantine approximation on \mathbf{G}_m* , Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 20 (1993), pp. 61–89.
- [18] E. BOMBIERI, *The equivariant Thue–Siegel method*, in Arithmetic geometry (Cortona, 1994), vol. XXXVII of Sympos. Math., Cambridge Univ. Press, Cambridge, 1997, pp. 70–86.
- [19] E. BOMBIERI, *Diophantine equations in low dimension*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., (2000), pp. 11–29. Mathematics towards the third millennium (Rome, 1999).
- [20] E. BOMBIERI, *Forty years of effective results in Diophantine theory*, in A panorama of number theory or the view from Baker’s garden (Zürich, 1999), Cambridge Univ. Press, Cambridge, 2002, pp. 194–213.
- [21] E. BOMBIERI AND P. B. COHEN, *Effective Diophantine approximation on \mathbf{G}_m . II*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 24 (1997), pp. 205–225.
- [22] E. BOMBIERI AND P. B. COHEN, *An elementary approach to effective Diophantine approximation on \mathbf{G}_m* , in Number theory and algebraic geometry, vol. 303 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 2003, pp. 41–62.
- [23] E. BOMBIERI AND W. GUBLER, *Heights in Diophantine Geometry*, no. 4 in Cambridge New Mathematical Monographs, Cambridge University Press, 2006.
- [24] E. BOMBIERI, A. J. VAN DER POORTEN, AND J. D. VAALER, *Effective measures of irrationality for cubic extensions of number fields*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 23 (1996), pp. 211–248.
- [25] J.-B. BOST, *Algebraic leaves of algebraic foliations over number fields*, Publ. Math. Inst. Hautes Études Sci., (2001), pp. 161–221.

- [26] J.-B. BOST AND F. CHARLES, *Quasi-projective and formal-analytic arithmetic surfaces*, 2022. arXiv:2206.14242v2.
- [27] F. BROWN, *Notes on motivic periods*, Commun. Number Theory Phys., 11 (2017), pp. 557–655.
- [28] F. BROWN AND W. ZUDILIN, *On cellular rational approximations to $\zeta(5)$* , 2022. arXiv:2210.03391v2.
- [29] Y. BUGEAUD, *Bornes effectives pour les solutions des équations en S -unités et des équations de Thue–Mahler*, J. Number Theory, 71 (1998), pp. 227–244.
- [30] Y. BUGEAUD, *Linear Forms in Logarithms and Applications*, vol. 28 of IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society (EMS), Zürich, 2018.
- [31] Y. BUGEAUD AND M. LAURENT, *Minoration effective de la distance p -adique entre puissances de nombres algébriques*, J. Number Theory, 61 (1996), pp. 311–342.
- [32] K. BUZZARD, *Analytic continuation of overconvergent eigenforms*, J. Amer. Math. Soc., 16 (2003), pp. 29–55.
- [33] F. CALEGARI, *Irrationality of certain p -adic periods for small p* , Int. Math. Res. Not., (2005), pp. 1235–1249.
- [34] F. CALEGARI, V. DIMITROV, AND Y. TANG, *The linear independence of 1, $\zeta(2)$, and $L(2, \chi_{-3})$* , 2024. arXiv:2408.15403.
- [35] F. CALEGARI, V. DIMITROV, AND Y. TANG, *The unbounded denominators conjecture*, J. Amer. Math. Soc., 38 (2025), pp. 627–702.
- [36] K. C. CHIM, *Lower bounds for linear forms in two p -adic logarithms*, J. Number Theory, 266 (2025), pp. 295–349.
- [37] D. V. CHUDNOVSKY AND G. V. CHUDNOVSKY, *Applications of Padé approximations to the Grothendieck conjecture on linear differential equations*, in Number theory (New York, 1983–84), vol. 1135 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 52–100.
- [38] G. V. CHUDNOVSKY, *On the method of Thue–Siegel*, Ann. of Math. (2), 117 (1983), pp. 325–382.
- [39] P. B. COHEN, *Perspectives de l’approximation diophantienne et de la transcendance*, 2003. Rankin memorial issues.
- [40] R. COLEMAN AND B. MAZUR, *The eigencurve*, in Galois representations in arithmetic algebraic geometry (Durham, 1996), vol. 254 of London Math. Soc. Lecture Note Ser., CUP, Cambridge, 1998, pp. 1–113.
- [41] P. DELIGNE, *Intégration sur un cycle évanescent*, Invent. Math., 76 (1984), pp. 129–143.
- [42] J.-H. EVERTSE AND K. GYÖRY, *Unit equations in Diophantine number theory*, vol. 146 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2015.
- [43] N. I. FELDMAN, *An effective power sharpening of a theorem of Liouville*, Izv. Akad. Nauk SSSR Ser. Mat., 35 (1971), pp. 973–990.
- [44] H. FURSTENBERG, *Algebraic functions over finite fields*, J. Algebra, 7 (1967), pp. 271–277.
- [45] H. FURUSHO, *p -adic multiple zeta values. I. p -adic multiple polylogarithms and the p -adic KZ equation*, Invent. Math., 155 (2004), pp. 253–286.
- [46] N. GOUILLON, *Explicit lower bounds for linear forms in two logarithms*, J. Théor. Nombres Bordeaux, 18 (2006), pp. 125–146.
- [47] K. IWASAWA, *Lectures on p -adic L -functions*, Annals of Mathematics Studies, No. 74, Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1972.

- [48] C. G. J. JACOBI, *Untersuchungen über die Differentialgleichung der hypergeometrischen Reihe*, J. Reine Angew. Math., 56 (1859), pp. 149–165.
- [49] M. KONTSEVICH AND D. ZAGIER, *Periods*, in Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 771–808.
- [50] L. LAI, J. SPRANG, AND W. ZUDILIN, *A note on the irrationality of $\zeta_2(5)$* , 2025. arXiv:2505.05005.
- [51] M. LAURENT, *Linear forms in two logarithms and interpolation determinants*, Acta Arith., 66 (1994), pp. 181–199.
- [52] M. LAURENT, M. MIGNOTTE, AND Y. NESTERENKO, *Formes linéaires en deux logarithmes et déterminants d’interpolation*, J. Number Theory, 55 (1995), pp. 285–321.
- [53] K. MAHLER, *Ein Beweis des Thue–Siegelschen Satzes über die Approximation algebraischer Zahlen für binomische Gleichungen*, Math. Ann., 105 (1931), pp. 267–276.
- [54] A. N. PARSHIN AND I. R. SHAFAREVICH, *Number theory. IV*, vol. 44 of Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 1998. Transcendental Numbers, by Naum I. Fel’dman and Yuri V. Nesterenko, Translation by N. Koblitz.
- [55] J.-P. SERRE, *Quelques propriétés des groupes algébriques commutatifs*, Astérisque, 69–70 (1979), pp. 191–202.
- [56] C. L. SIEGEL, *Approximation algebraischer Zahlen*, Math. Z., 10 (1921), pp. 173–213.
- [57] C. L. SIEGEL, *Die Gleichung $ax^n - by^n = c$* , Math. Ann., 114 (1937), pp. 57–68.
- [58] H. M. STARK, *Further advances in the theory of linear forms in logarithms*, in Diophantine approximation and its applications (Proc. Conf., Washington, D.C., 1972), Academic Press, 1973, pp. 255–293.
- [59] A. VAN DER POORTEN, *A proof that Euler missed...Apéry’s proof of the irrationality of $\zeta(3)$* , Math. Intelligencer, 1 (1978/79), pp. 195–203. An informal report.
- [60] C. VIOLA, *On Dyson’s lemma*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 12 (1985), pp. 105–135.
- [61] M. WALDSCHMIDT, *Minorations de combinaisons linéaires de logarithmes de nombres algébriques*, Canad. J. Math., 45 (1993), pp. 176–224.
- [62] M. WALDSCHMIDT, *Diophantine approximation on linear algebraic groups: Transcendence properties of the exponential function in several variables*, vol. 326 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, 2000.
- [63] G. YAMASHITA, *Bounds for the dimensions of p -adic multiple L -value spaces*, Doc. Math., (2010), pp. 687–723.
- [64] G. YAMASHITA, *p -adic multiple zeta values, p -adic multiple L -values, and motivic Galois groups*, in Galois–Teichmüller theory and arithmetic geometry, vol. 63 of Adv. Stud. Pure Math., Math. Soc. Japan, Tokyo, 2012, pp. 629–658.
- [65] K. YU, *Linear forms in p -adic logarithms. II*, Compositio Math., 74 (1990), pp. 15–113.
- [66] K. YU, *p -adic logarithmic forms and group varieties. I*, J. Reine Angew. Math., 502 (1998), pp. 29–92.
- [67] K. YU, *p -adic logarithmic forms and group varieties. II*, Acta Arith., 89 (1999), pp. 337–378.
- [68] K. YU, *p -adic logarithmic forms and group varieties. III*, Forum Math., 19 (2007), pp. 187–280.
- [69] U. ZANNIER, ed., *On some applications of Diophantine approximations*, vol. 2 of Quaderni/Monographs, 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.