

An Arithmetic Characterization of Some Algebraic Functions and a New Proof of an Algebraicity Prediction by Golyshev

Alin Bostan*

Inria, Université Paris-Saclay, 1 rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

*Correspondence to be sent to: e-mail: alin.bostan@inria.fr

Communicated by Prof. Umberto Zannier

We provide a new arithmetic characterization for the sequence of coefficients of algebraic power series $f(t)$ having the property that the differential equation $y'(t) = f(t)y(t)$ has algebraic solutions only. This extends a recent result by Delaygue and Rivoal, and also provides a new and shorter proof of an algebraicity result predicted by Golyshev.

1 Introduction

Let $\ell \in \mathbb{N}$ and assume that $\mathbf{c} = (c_1, \dots, c_\ell) \in \mathbb{N}^\ell$ and $\mathbf{d} = (d_1, \dots, d_{\ell+1}) \in \mathbb{N}^{\ell+1}$ are tuples of nonnegative integers such that $\sum_i c_i = \sum_j d_j$ and such that the sequence $(a_n)_{n \geq 0}$ defined by

$$a_n := \frac{(c_1 n)! \cdots (c_\ell n)!}{(d_1 n)! \cdots (d_{\ell+1} n)!} \quad (1)$$

is an integer sequence. The generating function

$$F(t) := \sum_{n=0}^{\infty} a_n t^n$$

is then hypergeometric and belongs to $\mathbb{Z}[[t]]$. Building on the “interlacing criterion” of Beukers and Heckman [4], Rodriguez-Villegas proved in [17] that F is algebraic over $\mathbb{Q}(t)$, that is, there exists a polynomial $P \in \mathbb{Q}[x, y] \setminus \{0\}$ such that $P(t, F(t)) = 0$. The simplest example is

$$F(t) = \sum_{n=0}^{\infty} \binom{cn}{dn} t^n, \quad \text{with } \ell = 1, c_1 = c, d_1 = d, d_2 = c - d,$$

where $c \geq d \geq 0$. More interesting examples (see also [5]) are given by the power series

$$\sum_{n=0}^{\infty} \frac{(6n)! n!}{(3n)! (2n)!^2} t^n, \quad \sum_{n=0}^{\infty} \frac{(10n)! n!}{(5n)! (4n)! (2n)!} t^n \quad \text{and} \quad \sum_{n=0}^{\infty} \frac{(30n)! n!}{(15n)! (10n)! (6n)!} t^n.$$

By Rodriguez-Villegas' result, these three power series are all algebraic. It can be proved (for instance by using a computer algebra system) that the algebraicity degrees of the first two ones are 6 and 30, while the degree of the third one is 483 840 according to [17, 18].

As accounted in Zagier's beautiful survey [20, pages 757–759], Golyshev predicted (based on “motivic arguments”) that the power series

$$Q(t) := \exp\left(\int \frac{F(t)}{t} dt\right) = t \exp\left(\sum_{n>0} a_n \frac{t^n}{n}\right) \quad (2)$$

is also an algebraic function. This “prediction” has been recently proved by Delaygue and Rivoal in [10]. Their proof consists of two steps:

(i) the sequence $(a_n)_{n \geq 0}$ defined in (1) has the *Gauss property*; this means that, for almost all prime numbers p , the Gauss congruences hold: $\frac{1}{n} \cdot (a_{np} - a_n) \in p\mathbb{Z}_{(p)}$ for all $n \geq 1$. (Here and in all the text, we denote by $\mathbb{Z}_{(p)}$ the localization of $\mathbb{Z}$ at the prime ideal (p) , that is the subring of $\mathbb{Q}$ consisting of rational numbers with denominator coprime to p .);

(ii) clearly, $y(t) := Q(t)/t$ satisfies $y'(t) = f(t)y(t)$, where $f(t) := \sum_{n>0} a_n t^{n-1}$; from this deduce that, as $(a_n)_{n \geq 0}$ has the Gauss property, $y(t)$ must be algebraic (and hence $Q(t)$ too).

Both steps are nontrivial. Step (i) is based on [9, Lemme 10], itself based on quite subtle p -adic congruences for Morita's p -adic Gamma functions. Step (ii) is based on the one hand on a consequence of Theorem 8.1 in [7] saying that if $g \in \mathbb{Z}[[\frac{t}{\lambda}]]$ for some $\lambda \in \mathbb{N} \setminus \{0\}$ is such that g'/g is algebraic, then g itself is algebraic; on the other hand, it is based on the classical Dieudonné–Dwork lemma [16, page 392] as well as on a refined p -adic analysis [10, §3.2.1] showing that $Q(t)$ defined in (2) belongs to $\mathbb{Z}[[\frac{t}{\lambda}]]$ for some $\lambda \in \mathbb{N} \setminus \{0\}$.

2 A New Proof of Golyshev's Prediction

In this note we provide a similar, but shorter, proof of the algebraicity of $Q(t)$ in (2). We directly use [7, Theorem 8.1], which gives a positive answer to Grothendieck's conjecture for linear differential equations of order 1 with algebraic coefficients. (Another proof of Thm. 8.1 in [7] is given in [2, Chap. VIII, Exs. 5 and 6]. The result was generalized by André to linear differential equations whose differential Galois groups have a solvable neutral component in [3, Cor. 4.3.5], see also [1, Thm. 2] and [8, Thm. 3.5]. It was further extended by Bost [6, Thm. 2.9] who proved a generalized version of the Grothendieck–Katz conjecture for any algebraic group, not necessarily linear, over an number field, whose neutral component is solvable. To our knowledge, all these proofs ultimately rely on algebraicity criteria generalizing the classical Borel–Dwork rationality criteria [11, Thms. 2 and 3], originally due to the Chudnovsky brothers [7, Thm. 5.9] and further extended by André and Bost in [2, Chap. VIII, Thms. 1.2 and 1.4], [3, Thm. 5.4.3 and Cor. 5.4.5], [6, Thms. 2.1 and 3.4], see also [8, Thms. 4.3 and 6.2].) Compared to the proof in [10] sketched above, we do not need Gauss congruences and we avoid the use of [9, Lemme 10], of the Dieudonné–Dwork lemma, and of the p -adic argument in [10, §3.2.1].

Instead, we prove in a direct and easy way that the sequence $(a_n)_{n \geq 0}$ satisfies the weaker congruences

$$a_{np} - a_n \in p\mathbb{Z} \quad \text{for all } n \geq 1,$$

and then conclude by using an explicit computation of the p -curvatures of $y'(t) = f(t)y(t)$.

Theorem 2.1. Let $(a_n)_{n \geq 1}$ be an integer sequence and assume that $f(t) := \sum_{n>0} a_n t^{n-1}$ is algebraic over $\mathbb{Q}(t)$. Then, all solutions of the differential equation $y'(t) = f(t)y(t)$ are algebraic if and only if, for almost all prime numbers p , the following infinite system of congruences holds

$$a_{np} \equiv a_n \pmod{p} \quad \text{for all } n \geq 1. \quad (3)$$

Proof. By [7, Theorem 8.1], all solutions of $y'(t) = f(t)y(t)$ are algebraic if and only if for almost all primes p , its p -curvatures are zero. To finish the proof, it is enough to show that, for any prime p , the vanishing of the p -curvature is equivalent to congruences (3). It is known since Jacobson's paper [14] that the p -curvature of $y'(t) = f(t)y(t)$ is equal (up to a sign) to $f^{(p-1)}(t) + f(t)^p$ modulo p . Therefore,

its vanishing is equivalent to $f(t)^p = -f^{(p-1)}(t) \pmod{p}$. To conclude, it is enough to remark that $(f^{(p-1)}(t))^{1/p} = -\sum_{n \geq 1} a_n t^{n-1}$. Indeed,

$$f^{(p-1)}(t) = \sum_{\ell \geq p} a_\ell (\ell-1)(\ell-2) \cdots (\ell-(p-1)) t^{\ell-p} = -\sum_{n \geq 1} a_{np} t^{(n-1)p} \pmod{p},$$

since the product $(\ell-1)(\ell-2) \cdots (\ell-(p-1))$ is nonzero modulo p if and only if p divides ℓ , in which case it is equal to -1 modulo p , by Wilson's theorem. ■

Theorem 2.2. Assume that $(a_n)_{n \geq 1}$ defined by (1) is an integer sequence with $\sum_i c_i = \sum_j d_j$. Then, for all prime numbers p , the sequence $(a_n)_{n \geq 1}$ satisfies the system of congruences (3).

Proof. We start by observing that for any prime p and any $m \in \mathbb{N}$, there exists $\alpha_m \in \mathbb{Z}$ with

$$(mp)! = m! (-p)^m (1 + p\alpha_m).$$

This is again a simple consequence of Wilson's theorem: indeed, $(mp)! = (1 \cdots (p-1)) \cdot p \cdot ((p+1) \cdots (2p-1)) \cdot (2p) \cdots ((m-1)p+1) \cdots (mp-1) \cdot (mp)$ is equal to $p \cdot (2p) \cdots (mp) \cdot (ps_1-1) \cdots (ps_m-1)$ for some $s_i \in \mathbb{Z}$, which rewrites $m! \cdot (-p)^m \cdot (1 + p\alpha_m)$ for some $\alpha_m \in \mathbb{Z}$.

It follows that there exist integer numbers $\alpha_{c,n}$ and $\alpha_{d,n}$ such that

$$a_{np} = \frac{\prod_{i=1}^{\ell} (c_i n p)!}{\prod_{j=1}^{\ell+1} (d_j n p)!} = \frac{\prod_{i=1}^{\ell} (c_i n)! (-p)^{c_i n} (1 + p\alpha_{c,n})}{\prod_{j=1}^{\ell+1} (d_j n)! (-p)^{d_j n} (1 + p\alpha_{d,n})} = a_n (-p)^{(\sum_i c_i - \sum_j d_j)n} \frac{\prod_{i=1}^{\ell} (1 + p\alpha_{c,n})}{\prod_{j=1}^{\ell+1} (1 + p\alpha_{d,n})}.$$

Since $\sum_i c_i = \sum_j d_j$ and the last products belong to $1 + p\mathbb{Z}_{(p)}$, the conclusion follows. ■

By combining Theorem 2.1 and Theorem 2.2, we deduce that Golyshev's prediction holds.

Corollary 2.3. Let $F(t)$ be an algebraic factorial generalized hypergeometric function, that is, a power series of the form

$$F(t) := \sum_{n=0}^{\infty} \frac{(c_1 n)! \cdots (c_\ell n)!}{(d_1 n)! \cdots (d_{\ell+1} n)!} t^n \quad (c_i, d_j \in \mathbb{N})$$

assumed to be algebraic over $\mathbb{Q}(t)$. Then, $\exp\left(\int \frac{F(t)}{t} dt\right)$ is also algebraic over $\mathbb{Q}(t)$.

Proof. By (the easy direction of) [17, Theorem 1], the algebraicity of F implies that its sequence of coefficients (a_n) is an integer sequence, that $\kappa = \ell + 1$ and $\sum_i c_i = \sum_j d_j$. In other words, (a_n) is of the form (1) and satisfies the assumptions of Theorem 2.2. By Theorem 2.2, (a_n) satisfies the system of congruences (3) for all prime numbers p . On the other hand, $f(t) := (F-1)/t = \sum_{n \geq 0} a_n t^{n-1}$ is algebraic, as F is algebraic. Now, $y := \exp\left(\int \frac{F(t)}{t} dt\right)/t$ satisfies the differential equation $y'(t) = f(t)y(t)$. The conclusion follows from Theorem 2.1. ■

We now state an alternative form of Golyshev's prediction.

Corollary 2.4. Let $\ell \in \mathbb{N}$ and assume that $(c_1, \dots, c_\ell) \in \mathbb{N}^\ell$ and $(d_1, \dots, d_{\ell+1}) \in \mathbb{N}^{\ell+1}$ are such that $\sum_i c_i = \sum_j d_j$ and such that the sequence $(a_n)_{n \geq 0}$ defined by

$$a_n := \frac{(c_1 n)! \cdots (c_\ell n)!}{(d_1 n)! \cdots (d_{\ell+1} n)!}$$

is an integer sequence. Then the power series

$$\sum_{n=0}^{\infty} a_n t^n \quad \text{and} \quad \exp\left(\sum_{n=1}^{\infty} a_n \frac{t^n}{n}\right)$$

are algebraic over $\mathbb{Q}(t)$.

Proof. By (the difficult direction of) [17, Theorem 1], $\sum_{n=0}^{\infty} a_n t^n$ is algebraic. The proof heavily relies on the interlacing criterion due to Beukers and Heckman [4]. The algebraicity of $\exp\left(\sum_{n=1}^{\infty} a_n \frac{t^n}{n}\right)$ can then be proved as in the second part of the proof of Corollary 2.3. ■

3 Main Result

We now prove a more general version of Theorem 2.1. This is the main result of this note. Not only we relax the integrality assumption on the sequence (a_n) , but we prove the equivalence of the system of congruences (3) for almost all primes p with the Gauss congruences considered in [10] for almost all primes p .

Theorem 3.1. Let $(a_n)_{n \geq 1}$ be a sequence of rational numbers and assume that $f(t) := \sum_{n \geq 0} a_n t^{n-1}$ is algebraic over $\mathbb{Q}(t)$. Then, the following assertions are equivalent:

- (a) all solutions of the differential equation $y'(t) = f(t)y(t)$ are algebraic over $\mathbb{Q}(t)$;
- (b) for almost all prime numbers p , we have that $a_{np} - a_n \in p\mathbb{Z}_{(p)}$ for all $n \geq 1$;
- (c) for almost all prime numbers p , we have that $a_{np} - a_n \in n p \mathbb{Z}_{(p)}$ for all $n \geq 1$.

Condition (c) is called the *Gauss property* for the sequence (a_n) (or, equivalently, for the power series $tf(t)$).

Note that the same statement holds true if $f(t)$ is assumed to belong more generally to $\mathbb{Q}((t))$, and to be algebraic over $\mathbb{Q}(t)$. Indeed, with this assumption one can show as in [10, §3.2] that if $y'(t) = f(t)y(t)$ has only algebraic solutions, then $tf(t)$ belongs to $\mathbb{Q}[[t]]$; similarly, if $(a_n)_{n \geq b}$ ($b \in \mathbb{Z}$) satisfies the Gauss property, then $a_n = 0$ for $n < 0$; and the same holds for the congruences in condition (b). In other words, proving Theorem 3.1 with the assumption $f \in \mathbb{Q}((t))$ reduces easily to the case $\mathbb{Q}[[t]]$ treated in Theorem 3.1.

Proof. Obviously, (c) implies (b), and Eisenstein's theorem together with the Dieudonné–Dwork criterion show that (a) implies (c) (see [10]). It remains to prove that (b) implies (a). This was proved in our Theorem 2.1 under the stronger assumption that $f \in \mathbb{Z}[[t]]$. As f is algebraic, Eisenstein's theorem implies that there exists $c \in \mathbb{N} \setminus \{0\}$ such that $g(t) := cf(tc)$ belongs to $\mathbb{Z}[[t]]$. In other words, the sequence $b_n := c^{n+1}a_n$ is an integer sequence. We shall prove that $b_{np} \equiv b_n \pmod{p}$ for almost all primes p and for all $n \geq 1$. Let $B \geq 1$ be such that $a_{np} - a_n \in p\mathbb{Z}_{(p)}$ for all $n \geq 1$ and all primes $p \geq B$. Let $B' \geq B$ be such that all prime divisors of c are less than B' . Then, for all primes $p \geq B'$ we have that $b_{np} - b_n = c^{np+1}a_{np} - c^{n+1}a_n = c(c^{np} - c^n)a_{np} + c^{n+1}(a_{np} - a_n)$ belongs to $p\mathbb{Z}_{(p)}$, because p divides the integer $c^{np} - c^n$, while $a_n \in \mathbb{Z}[\frac{1}{c}] \subset \mathbb{Z}_{(p)}$ and $a_{np} - a_n \in p\mathbb{Z}_{(p)}$. Thus, $b_{np} - b_n \in \mathbb{Z} \cap p\mathbb{Z}_{(p)} = p\mathbb{Z}$ for all primes $p \geq B'$ and for all $n \geq 1$. By Theorem 2.1, all solutions of the differential equation $y'(t) = g(t)y(t)$ are algebraic. As $g(t) = cf(tc)$, this implies that all solutions of the differential equation $y'(t) = f(t)y(t)$ are algebraic as well. ■

4 Conclusion and Further Questions

A surprising consequence of our results is the equivalence between conditions (b) and (c) in Theorem 3.1 when each holds for *almost all primes* p . This equivalence is a priori unexpected, and it is in fact the main point in our simpler proof in Section 2: to prove Golyshev's prediction, instead of checking the Gauss property for the factorial ratios sequences, it is enough to check the more relaxed congruences (3). Let us remark that (b) and (c) are in general not equivalent for one prime p . On the other hand, the equivalence between (a) and (b) does not hold if one replaces “for almost all primes p ” by “for an infinite number of primes p ”: for instance, when $f(t) = 1/(t^2 + 1)$, (a) does not hold, but (b) and (c) hold for all primes p equal to 1 modulo 4.

Independently of the algebraicity context of Theorem 3.1, it is legitimate to ask for which integer (or, globally bounded) sequences (a_n) does the equivalence between (b) and (c) hold? For instance, does it hold for (some of) the integer factorial ratios sequences of the form $a_n = \prod_{i=1}^{\ell} (c_i n)! / \prod_{j=1}^k (d_j n)!$ studied by Soundararajan in [19]? (See also [12].)

Another natural question is whether there exists a “finite version” of Theorems 2.1 and 3.1. By this, we mean: given the algebraic function $f(t)$, can one find effective bounds $B_1(f)$ and $B_2(f)$ such that one can replace conditions (b) and (c) in Theorem 3.1 by their finite versions:

- (b') $a_{np} - a_n \in p\mathbb{Z}_{(p)}$ for all $n \geq 1$ and all primes $p \in [B_1(f), B_2(f)]$;
 (c') $a_{np} - a_n \in n p \mathbb{Z}_{(p)}$ for all $n \geq 1$ and all primes $p \in [B_1(f), B_2(f)]$.

As the proof of Theorem 3.1 is ultimately based on Chudnovskys' Theorem 8.1 in [7], this question amounts to asking whether there exists a finite version of that theorem.

Another natural question is whether Corollary 2.3 and Corollary 2.4 can be proved without relying on Chudnovskys' theorem at all. For instance, can it be deduced using the notion of p^i -curvatures studied in the recent preprints [13] and [15]?

Finally, one may wonder whether more explicit structural information can be obtained in the conclusion of Corollary 2.3, for example, concerning the algebraicity degrees. For instance, can they be expressed in terms of the numbers c_i and d_j ? According to Zagier [20, page 757], Golyshev had also predicted that, under the assumptions of Corollary 2.3, the power series $\exp(\sum_{n=1}^{\infty} a_n t^n/n)$ has the same algebraicity degree as $\sum_{n=0}^{\infty} a_n t^n$, and moreover that it is an algebraic unit over $\mathbb{Z}[1/t]$. To our knowledge, all these are open questions.

Funding

This work was supported by the French grant DeRerumNatura [ANR-19-CE40-0018] and the French–Austrian project EAGLES [ANR-22-CE91-0007 and FWF I6130-N].

Acknowledgments

Warm thanks go to Xavier Caruso for enlightening discussions, to Éric Delaygue and Sergey Yurkevich for their very careful reading of preliminary versions of this text, and to Gilles Christol, Vasily Golyshev, Maxim Kontsevich, Hossein Movasati, Umberto Zannier, Wadim Zudilin, and the two anonymous reviewers for their positive feedback.

References

1. André, Y. “Quatre descriptions des groupes de Galois différentiels.” In *Séminaire d'algèbre Paul Dubreil et Marie-Paule Malliavin (Paris, 1986)*. Lecture Notes in Math., vol. 1296, 28–41. Springer, Berlin, 1987, <https://doi.org/10.1007/BFb0078522>.
2. André, Y. “G-functions and geometry.” In *Aspects of Mathematics*, E13, xii+229 pp. ISBN: 3-528-06317-3. Braunschweig: Friedr. Vieweg & Sohn, 1989, <https://doi.org/10.1007/978-3-663-14108-2>.
3. André, Y. “Sur la conjecture des p -courbures de Grothendieck–Katz et un problème de Dwork.” In *Geometric Aspects of Dwork Theory*. Vol. I, II. 55–112. Berlin: Walter de Gruyter, 2004, <https://doi.org/10.1515/9783110198133.1.55>.
4. Beukers, F. and G. Heckman. “Monodromy for the hypergeometric function ${}_nF_{n-1}$.” *Invent. Math.* **95**, no. 2 (1989): 325–54. <https://doi.org/10.1007/BF01393900>.
5. Bober, J. W. “Factorial ratios, hypergeometric series, and a family of step functions.” *J. Lond. Math. Soc.* (2) **79**, no. 2 (2009): 422–44. <https://doi.org/10.1112/jlms/jdn078>.
6. Bost, J.-B. “Algebraic leaves of algebraic foliations over number fields.” *Publ. Math. Inst. Hautes Études Sci.* **93** (2001): 161–221. <https://doi.org/10.1007/s10240-001-8191-3>.
7. Chudnovsky, D. V. and G. V. Chudnovsky. “Applications of Padé approximations to the Grothendieck conjecture on linear differential equations.” In *Number Theory (New York, 1983–84)*. LNMS, vol. 1135, 52–100. Springer, 1985. <https://doi.org/10.1007/BFb0074601>.
8. Chambert-Loir, A. Théorèmes d'algébricité en géométrie diophantienne (d'après J.-B. Bost, Y. André, D. & G. Chudnovsky). In *Séminaire Bourbaki*, vol. 2000/2001, talks 880–893. Astérisque vol. 282, talk no. 886, 175–209, 2002. http://www.numdam.org/item/SB_2000-2001__43__175_0/.
9. Delaygue, E. “Critère pour l'intégralité des coefficients de Taylor des applications miroir.” *J. Reine Angew. Math.* **662** (2012): 205–52. <https://doi.org/10.1515/CRELLE.2011.094>.
10. Delaygue, É. and T. Rivoal. “On Abel's problem and Gauss congruences.” *Int. Math. Res. Not. IMRN* **2024**, no. 5 (2024): 4301–27. <https://doi.org/10.1093/imrn/rnad229>.

11. Dwork, B. "On the rationality of the zeta function of an algebraic variety." *Am. J. Math.* **82**, no. 3 (1960): 631–48. <https://doi.org/10.2307/2372974>.
12. Dvornicich, R. and U. Zannier. "Fractional parts of linear polynomials and an application to hypergeometric functions." *J. Aust. Math. Soc.* **70**, no. 3 (2001): 401–24. <https://doi.org/10.1017/S144678870000241X>.
13. Fürnsinn, F., H. Hauser, and H. Kawanoue. "On Abel's problem about logarithmic integrals in positive characteristic." 2024, <https://arxiv.org/abs/2401.14154>.
14. Jacobson, N. "Abstract derivation and lie algebras." *Trans. Amer. Math. Soc.* **42**, no. 2 (1937): 206–24. <https://doi.org/10.1090/S0002-9947-1937-1501922-7>.
15. Movasati, H. "Grothendieck–Katz conjecture." 2024, <https://arxiv.org/abs/2403.11829>.
16. Robert, A., M. "A course in p -adic analysis." *Graduate Texts in Mathematics*, vol. 198. New York: Springer, 2000. xvi+438 pp. ISBN: 978-0-387-98669-2. <https://doi.org/10.1007/978-1-4757-3254-2>.
17. Rodriguez-Villegas, F. "Integral ratios of factorials and algebraic hypergeometric functions." *Oberwolfach Rep.* **32** (2005): 1813–6. https://publications.mfo.de/bitstream/handle/mfo/2906/OWR_2005_32.pdf.
18. Rodriguez-Villegas, F. "Mixed Hodge numbers and factorial ratios." 2019, <https://arxiv.org/abs/1907.02722>.
19. Soundararajan, K. "Integral factorial ratios." *Duke Math. J.* **171**, no. 3 (2022): 633–72. <https://doi.org/10.1215/00127094-2021-0017>.
20. Zagier, D. "The arithmetic and topology of differential equations." *European Congress of Mathematics*. 717–76. Zürich: Eur. Math. Soc., 2018. <https://doi.org/10.4171/176-1/33>.