

Non-Holonomicity of Sequences Defined via Elementary Functions

Jason P. Bell¹, Stefan Gerhold^{2*}, Martin Klazar^{3†}, and Florian Luca⁴

¹Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, B.C.
V5A 1S6, Canada
jpb@sfu.ca

²Christian Doppler Laboratory for Portfolio Risk Management, Vienna University of Technology,
Wiedner Hauptstraße 8/105-1 A-1040 Vienna, Austria
sgerhold@fam.tuwien.ac.at

³Department of Applied Mathematics (KAM) and Institute for Theoretic Computer Science
(ITI), Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25,
118 00 Praha, Czech Republic
klazar@kam.mff.cuni.cz

⁴Instituto de Matemáticas UNAM, C.P. 58 089 Campus Morelia, Michoacán, Mexico
fluca@matmor.unam.mx

Received May 5, 2006

AMS Subject Classification: 11B37, 11B83

Abstract. We present a new method for proving non-holonomicity of sequences, which is based on results about the number of zeros of elementary and of analytic functions. Our approach is applicable to sequences that are defined as the values of an elementary function at positive integral arguments. We generalize several recent results, e.g., non-holonomicity of the logarithmic sequence is extended to rational functions involving $\log n$. Moreover, we show that the sequence that arises from evaluating the Riemann zeta function at an increasing integer sequence with bounded gap lengths is not holonomic.

Keywords: holonomic sequences, fewnomials, meromorphic functions

1. Introduction

A sequence of complex numbers $(a_n)_{n \geq 1}$ is called *holonomic* (P-recursive) if there exist polynomials $p_k(x) \in \mathbb{C}[x]$, $k = 0, 1, \dots, d$, which are not all zero, such that

$$p_0(n)a_n + p_1(n)a_{n+1} + \dots + p_d(n)a_{n+d} = 0 \quad (1.1)$$

* Supported in part by the Christian Doppler Research Association (CDG). S. Gerhold gratefully acknowledges a fruitful collaboration and continued support by Bank Austria and ÖBFA through CDG. Supported in part by the FWF grant F1305.

† ITI is supported as project 1M0021620808 by Ministry of Education of the Czech Republic.

holds for every integer $n \geq 1$ (equivalently, for every integer $n \geq n_0$). Their continuous counterpart are holonomic (D -finite) functions $g(z)$, which satisfy linear differential equations

$$p_0(z)g(z) + p_1(z)g'(z) + \cdots + p_d(z)g^{(d)}(z) = 0$$

with polynomial coefficients. For more information on holonomic sequences and functions, especially in combinatorial enumeration, see Stanley [22]. For the sake of clarity, in this paper we will use the word holonomic for sequences and the word D -finite for functions. Sequences that satisfy (1.1) with $d = 1$ are called hypergeometric. An important property for our purposes is that the sum and the product of two holonomic sequences are holonomic.

The role of the holonomic sequences in the set of all complex sequences is reminiscent of the role of the algebraic numbers in the set of all complex numbers. Just as complex numbers are usually not algebraic, unless they are by design, a sequence that is not obviously holonomic is usually not holonomic. In both situations it can be a challenge, though, to come up with a proof. Flajolet, Gerhold, and Salvy [3] gave an exhaustive survey of known non-holonomicity results. In the present article we shall be interested in the following:

Question. If $f: \Omega \rightarrow \mathbb{C}$ is a “nice” function that is defined on a domain $\Omega \subset \mathbb{C}$ containing all natural numbers $\mathbb{N} = \{1, 2, 3, \dots\}$, when is it the case that the sequence of values of f on \mathbb{N} , $(f(n))_{n \geq 1}$, is non-holonomic?

We seek general results and criteria yielding non-holonomicity of $(f(n))_{n \geq 1}$ for f from large classes of functions \mathcal{F} , generalizing the results from Gerhold [6] and some of the results of Flajolet, Gerhold, and Salvy [3]. In the first paper, $(\log n)_{n \geq 1}$ was proved conditionally to be non-holonomic, and in the second paper an unconditional proof was given by an asymptotic machinery. A simple proof was given also by Klazar [10]. Here, we generalize the method from the latter paper and arrive at non-holonomicity results for many sequences for which closed-form expressions are available.

In the following section we explain our proving method. Section 3 is devoted to straightforward applications of our approach. In Section 4 we develop our arguments further in several directions to establish the non-holonomicity of some sequences involving the exponential and factorial functions. We proceed by presenting several unrelated results, e.g., about algebraic sequences and interlacement sequences in Section 5. Finally, we show in Section 6 that the sequence arising from evaluating the Riemann zeta function at an increasing sequence of integers with bounded gap lengths is not holonomic. The method we use there is not related to our main approach.

2. The Proving Method

Before presenting our approach we briefly discuss the method of Flajolet, Gerhold, and Salvy [3]. They use the equivalence saying that $(a_n)_{n \geq 1}$ is holonomic if and only if the ordinary generating function $g(z) = \sum_n a_n z^n$ is D -finite [22]. The asymptotic behavior of a holonomic function near a singularity is constrained by a structure theorem [3, Theorem 2], and Abelian theorems transfer the asymptotic behavior of a_n as $n \rightarrow \infty$ to the asymptotic behavior of $g(z)$ near the singularity [3, Theorem 3]. Thus, if $(a_n)_{n \geq 1}$ (or any transform of it obtained by holonomicity-preserving transformations)

has an asymptotic behaviour that is transferred to “forbidden” asymptotics of the generating function, then $(a_n)_{n \geq 1}$ is not holonomic. While requiring an initial asymptotic analysis of $(a_n)_{n \geq 1}$, which is a drawback, the real strength of this method is that this asymptotic behaviour is fully sufficient information, and a closed-form representation $a_n = f(n)$ is not needed to prove non-holonomicity. This has been demonstrated, e.g., on the sequence of primes [3]. In contrast, our method relies heavily on the explicit representation $a_n = f(n)$.

We associate with a function $f(z)$ and a $(d + 1)$ -tuple of complex polynomials $p_0(z), \dots, p_d(z)$ the function

$$F(z) = F(z; f, p_0, \dots, p_d) := \sum_{i=0}^d p_i(z) f(z+i).$$

If $(f(n))_{n \geq 1}$ is holonomic, then, for some not-all-zero polynomials p_0, \dots, p_d , the function $F(z)$ vanishes at $z = 1, 2, \dots$. If f is real-valued in the positive reals (or at least in \mathbb{N}), then $(f(n))_{n \geq 1}$ is a sequence of real numbers, and by an algebraic argument [6, 12] the polynomials in (1.1) can be assumed without loss of generality to have real coefficients. This is important for the formulation of the following two properties for classes of complex functions \mathcal{F} : If all $f \in \mathcal{F}$ are real-valued for real arguments, we assume that the polynomials $p_0(z), \dots, p_d(z)$ have real coefficients.

Property (A) If $f \in \mathcal{F}$ is not identically zero and not all (real or complex, see above) polynomials $p_0(z), \dots, p_d(z)$ are zero, then the function $F(z; f, p_0, \dots, p_d)$ is not identically zero.

Property (B) If $F(z) = F(z; f, p_0, \dots, p_d)$, $f \in \mathcal{F}$, vanishes at all $z \in \mathbb{N}$, then $F(z)$ is identically zero.

It follows that any non-identically-zero function f from a class \mathcal{F} with properties (A) and (B) produces a non-holonomic sequence $(f(n))_{n \geq 1}$. (For some of the classes \mathcal{F} we are going to discuss, 1 has to be replaced by some other positive integer n_0 .)

Let us have a look at some (classes of) functions having property (A) and then at functions having property (B). Ultimately we want to have large classes of functions having simultaneously properties (A) and (B).

Condition (A1). Meromorphic functions and singularities. The class of functions f which are meromorphic for $\Re(z) > 0$ but have no extension that is meromorphic at $z = 0$ has property (A). Indeed, if p_0 is a non-zero polynomial and $F(z; f, p_0, \dots, p_d) \equiv 0$, then

$$f(z) = -p_0(z)^{-1} \sum_{i=1}^d p_i(z) f(z+i), \quad (2.1)$$

for $\Re(z) > 0$, which is impossible, because the right-hand side is meromorphic at $z = 0$ while the left-hand side is not. This criterion can be generalized in an obvious way by replacing 0 with any other point $z_0 \in \mathbb{C}$. Examples of such functions are $\log z$, $\log \log z$, z^α with $\alpha \in \mathbb{C} \setminus \mathbb{Z}$, $\exp(\sqrt{z})$, $\exp(\sqrt{\log z})$, $\sin(1/z)$, z^z , $\arctan z$, etc. Condition (A1) will be our main tool for establishing property (A). Note how the assumption that our sequences a_n are defined via functions $f(z)$ enables us to work with the asymptotic behaviour of $f(z)$ at points other than infinity.

Condition (A2). Logarithmic derivatives and shifts. Here we assume that $f'(z)/f(z) \in \mathbb{C}(z)$, the field of rational functions with complex coefficients. Notice that this implies that f is of the form

$$f(z) = \exp(r(z)) \prod_{i=1}^m (z - \alpha_i)^{c_i},$$

where $r(z)$ is a rational function, and $\alpha_i, c_i \in \mathbb{C}$. (Indeed, if $f'/f = q \in \mathbb{C}(z)$, then $f = C \exp(\int q)$ for some complex C , and the integral of a rational function can always be expressed as a rational function plus a linear combination of logarithms of linear functions.)

Proposition 2.1. *The class of functions $f(z)$ that are analytic in a neighborhood of $[1, \infty[$ and satisfy $f'(z)/f(z) \in \mathbb{C}(z)$ but $f(z+k)/f(z) \notin \mathbb{C}(z)$, $k \in \mathbb{N}$, has property (A).*

Proof. Suppose, on the contrary, that we have a nontrivial relation $F(z; f, p_0, \dots, p_d) \equiv 0$, with minimal d . Upon differentiating, this implies

$$0 = F'(z) = \sum_{i=0}^d \left(p'_i(z) + p_i(z) \frac{f'(z+i)}{f(z+i)} \right) f(z+i).$$

We now eliminate the summand $i = 0$ from the equations $F(z) = 0$ and $F'(z) = 0$. The coefficients $b_i(z)$ in

$$\sum_{i=1}^d b_i(z) f(z+i) := F'(z) p_0(z) - F(z) \left(p'_0(z) + p_0(z) \frac{f'(z)}{f(z)} \right) \in \mathbb{C}(z)$$

are in $\mathbb{C}(z)$, hence they vanish by the minimality of d . For $i = d$ this reads as

$$0 = p_0(z) \left(p'_d(z) + p_d(z) \frac{f'(z+d)}{f(z+d)} \right) - \left(p'_0(z) + p_0(z) \frac{f'(z)}{f(z)} \right) p_d(z).$$

Dividing by $p_0(z)p_d(z)$ and integrating, we obtain

$$f(z+d)/f(z) = C p_0(z)/p_d(z) \in \mathbb{C}(z)$$

for some constant C , a contradiction. Note that we have $p_0(z)p_d(z) \neq 0$ by the minimality of d . ■

Notice that the field $\mathbb{C}(z)$ in the proposition can be replaced by an arbitrary field $\mathbb{K} \supset \mathbb{C}(z)$ of functions meromorphic at infinity, provided that it is closed under differentiation and shift. For instance, this generalization can be applied to $f(z) = \text{expexp}(1/z)$ with $\mathbb{K} = \mathbb{C}(z, e^{1/z}, e^{1/(z+1)}, \dots)$.

We remark that we do not know of examples where condition (A2) works and condition (A1) fails. However, Proposition 2.1 can be viewed as a first step to address the question which D -finite functions $f(z)$ define holonomic sequences $(f(n))_{n \geq 1}$. The interplay of shift and derivative in the proof of Proposition 2.1 shows that this problem is not as contrived as it might seem.

Condition (A3). Growth conditions. If $f(z)$ is such that $|f(z+1)/f(z)|$ grows faster than polynomially as $|z| \rightarrow \infty$ then (2.1) cannot hold identically either. We will not use

this observation in what follows, since it is well-known that sequences that grow faster than a power of $n!$ cannot be holonomic.

We now present some sufficient conditions for property (B).

Condition (B1). Rational functions. The class of rational functions has property (B). This is trivial but by itself not very useful, since rational functions do not have property (A). Besides, rational functions clearly define holonomic sequences.

Condition (B2). Differentiation and Rolle's theorem. If $F(z)$ is a non-identically-zero real smooth function and has infinitely many real positive zeros, then by Rolle's theorem so have $F'(z)$ and all higher derivatives of $F(z)$. Thus, if for sufficiently large k we get a function $F^{(k)}(z)$ with only finitely many real positive zeros (e.g., a rational function, cf. condition (B1)), then $F(z)$ has only finitely many zeros as well. This argument shows that, for example, the class of functions that lie in $\mathbb{R}[x, \log x]$ and are defined in $[1, \infty[$ has property (B).

Condition (B3). Meromorphicity at infinity. The class of functions $f(z)$ meromorphic at ∞ has property (B). Indeed, the function $F(z) = F(z; f, p_0, \dots, p_d)$ is then meromorphic at ∞ as well and cannot have infinitely many zeros in \mathbb{N} , unless it is identically zero. This applies, e.g., to $\sin(1/z)$, $\exp(1/z)$, etc.

Condition (B4). Classical results about zeros of analytic functions. Many authors have investigated the distribution of zeros of analytic functions. We just recall the following classical theorem [15]. Its gist is that the function $\sin\pi z$ is the “smallest” function analytic in the right half-plane that vanishes at $0, 1, 2, \dots$

Theorem 2.2. [Carlson] *Let $g(z)$ be a function analytic in $\Re(z) \geq 0$ that satisfies the growth conditions $g(z) = O(\exp(\alpha|z|))$ and $g(\pm ix) = O(\exp((\pi - \varepsilon)x))$, $x > 0$, for some positive constants α and ε . If $g(z)$ vanishes at $z = 0, 1, 2, \dots$, then $g(z)$ is identically zero.*

Corollary 2.3. *The class of functions $f(z)$ such that $g(z) = f(z + 1)$ satisfies the conditions of Theorem 2.2 has property (B).*

Condition (B5). Zeros of real elementary functions – results of Khovanskii. In [9], Khovanskii proved rather general results on the number of simultaneous zeros of multivariate real elementary functions. We will use his results only for univariate functions. Adopting Khovanskii's definition, we call a univariate function $g(x)$ *elementary* if it can be expressed as a composition of multivariate rational functions and the functions $\exp(x)$, $\log(x)$, $\sin(x)$, $\cos(x)$, $\tan x$, $\arcsin(x)$, $\arccos(x)$, and $\arctan(x)$. The domain of definition of $g(x)$ must be such that denominators do not vanish, arguments of \log are positive, arguments of \arcsin and \arccos lie in $[-1, 1]$, arguments of \tan lie in $[-\pi/2, \pi/2]$, and arguments of \sin and \cos are bounded. The latter assumption is crucial to recover the finiteness of the zero set. Then a special case of one of Khovanskii's main results [9, §1.6] says:

Theorem 2.4. [Khovanskii] *An elementary function $g(x)$ has only finitely many simple zeros in its domain of definition.*

Though apparently only a technical nuisance, the restriction to simple zeros makes this result somewhat unwieldy for our purposes. To overcome this difficulty, we impose the additional assumption that $f(x)$ be a quotient of D -finite functions. Then $F(x)$ is also a quotient of D -finite functions, and it is easy to see that the order of the zeros of $F(x)$ is therefore bounded. Thus, by repeated differentiation, we arrive at a function with only simple zeros, to which Theorem 2.4 can be applied.

Corollary 2.5. *The class of elementary functions that are defined on $[1, \infty[$ and are a quotient of D -finite functions has property (B).*

Let us briefly compare this to Carlson's theorem. Apart from the absence of a growth condition, Khovanskii's result does not require analyticity in a right half-plane. Thus, for instance, we need not worry about zeros of the denominator of $f(x)$ outside the positive reals. On the other hand, Khovanskii's result is about *real* functions only, and his restriction to simple zeros induced us to include the additional D -finiteness requirement.

3. First Results

Putting together the results from the preceding section immediately yields several non-holonomicity proofs without any additional work.

Theorem 3.1. *Let $f(z)$ be a function from $\mathbb{R}(z, \log z, \arctan z)$ that is analytic in an open set containing the positive reals. Then $(f(n))_{n \geq 1}$ is a holonomic sequence if and only if $f(z)$ is a rational function.*

Proof. The result is an immediate consequence of conditions (A1) and (B2) (Corollary 2.5). Indeed, if $f(z)$ depends on $\log z$, then $F(z) = F(z; f, p_0, \dots, p_d)$ has a logarithmic singularity at $z = 0$, and if it depends on $\arctan z$, then $F(z)$ has a logarithmic singularity at $z = i$. ■

We have excluded the entire functions \exp and \sin from the preceding theorem, since condition (A1) does not work for them in general. In Section 4 we will completely describe the rational functions of n and e^n that define holonomic sequences. As long as there is some \log or \arctan present, the argument from the proof of Theorem 3.1 goes through: For instance,

$$a_n = \frac{\log n + \sin \frac{1}{n+1}}{e^n + n^2 + 1}$$

does not define a holonomic sequence for this reason. As for compositions, functions like $\log \log z$ are not allowed in Theorem 3.1, since D -finite functions are not closed under composition, violating our D -finiteness assumption that takes care of multiple zeros of elementary functions. (The sequence $(\log \log n)_{n \geq 2}$ can be treated by the asymptotic method [3] for showing non-holonomicity.) Still, D -finite functions *are* closed under composition with algebraic functions. For example,

$$f(z) = \frac{e^{z^2} + 1}{e^{z^2} - 1}$$

is a quotient of D -finite functions, hence the assumptions of Corollary 2.5 are satisfied. The complex poles of $f(z)$ are at $z = \pm\sqrt{2k\pi}e^{i\pi/4}$, $k \in \mathbb{Z}$, which shows that condition (A1) can be applied, establishing the non-holonomicity of the corresponding sequence. Observe that Carlson's theorem is not (directly) applicable here, since $f(z)$ is not analytic in a right half-plane. Anyhow, let us now recapitulate what we can infer from Corollary 2.3.

Theorem 3.2. *Let $f(z)$ be a function analytic for $\Re(z) > 0$ that does not have an extension that is meromorphic at $z = 0$. If $f(z+1)$ satisfies the growth conditions in Theorem 2.2, then the sequence $(f(n))_{n \geq 1}$ is not holonomic.*

This gives another proof that $\log n$ is not holonomic, and also establishes non-holonomicity of $e^{\sqrt{n}}$, $e^{1/n}$, $\exp\exp(1/n)$, and n^α with $\alpha \in \mathbb{C} \setminus \mathbb{Z}$. Note that for the latter three we could have used condition (A2) instead of (A1) to establish that F cannot vanish identically. Also, Khovanskii's theorem can be used instead of Carlson's in all five cases. Observe that, to our knowledge, this is the first published proof that $e^{\sqrt{n}}$ and $\exp\exp(1/n)$ are not holonomic.

We have seen by now that our method works on several sequences without any effort. Take for example the sequence $e^{1/n}$. Proving its non-holonomicity by asymptotics requires a somewhat involved asymptotic analysis [7] of the generating function $\sum_{n \geq 1} e^{1/n} z^n$. But longer proofs can of course have merits too; this analysis has stimulated comprehensive further asymptotic investigations [4] of analytic functions of the form $\sum_{n \geq 1} a^{n^b} z^n$, which are of interest in their own right.

4. Factorial and Exponential Sequences

We now present some less obvious applications of the tools we have collected in Section 2. We start with some rather concrete sequences and give a few more general results, most notably about algebraic sequences, in the following section. Let us begin with the sequence $n!^\alpha$, $\alpha \in \mathbb{C} \setminus \mathbb{Z}$. It grows too fast to apply Carlson's theorem. Also, the function $\Gamma(z+1)^\alpha$ is not a quotient of D -finite functions, hence Corollary 2.5 is not applicable. We now show that this sequence can be knocked out by Carlson's theorem after a little bit of massage on the alleged recurrence.

Theorem 4.1. *Let u_1, \dots, u_s be distinct complex numbers, and let $\alpha_1, \dots, \alpha_s$ be complex numbers. Then the sequence $\Gamma(n - u_1)^{\alpha_1} \cdots \Gamma(n - u_s)^{\alpha_s}$ is holonomic if and only if $\alpha_1, \dots, \alpha_s$ are integers.*

In particular, powers of hypergeometric sequences with non-integral exponent are not holonomic, unless the exponent trivially cancels, such as in $(n!^2)^{1/2}$. This generalizes a known result [6, Theorem 1].

Proof. Suppose that the sequence is holonomic. Since hypergeometric sequences are holonomic, we may assume without loss of generality that no α_i is an integer. The sequence is defined for $n \geq n_0 := \lfloor \sigma \rfloor + 1$, where $\sigma := \max_i \Re(u_i)$. Suppose that $\sum_{k=0}^d p_k(z) f(z+k)$ vanishes for $z = n_0, n_0 + 1, \dots$, where the $p_k(z)$ are polynomials, $p_0(z)$ is not identically zero, and

$$f(z) := \Gamma(z - u_1)^{\alpha_1} \cdots \Gamma(z - u_s)^{\alpha_s}.$$

By the recurrence of $\Gamma(z)$ and the fact that $\Gamma(z)$ has no zeros, this implies

$$\sum_{k=0}^d p_k(z) \prod_{i=1}^s \left((z - u_i)^{\bar{k}} \right)^{\alpha_i} = 0, \quad z = n_0, n_0 + 1, \dots, \quad (4.1)$$

where $z^{\bar{k}}$ denotes the rising factorial

$$z^{\bar{k}} = z(z+1)\cdots(z+k-1).$$

Replacing z by $z+n_0$ and applying Theorem 2.2, we find that the left-hand side of (4.1) vanishes for $\Re(z) \geq n_0$, hence by the identity theorem it vanishes in \mathbb{C} slit along s rays from u_i to $-\infty$. We thus have

$$\begin{aligned} 0 &= p_0(z) + (z - u_1)^{\alpha_1} \cdots (z - u_s)^{\alpha_s} \sum_{k=1}^d p_k(z) \prod_{i=1}^s \left((z - u_i + 1)^{\bar{k}-1} \right)^{\alpha_i} \\ &=: p_0(z) + (z - u_1)^{\alpha_1} G(z). \end{aligned}$$

If we assume without loss of generality that the u_i are in descending order with respect to the size of their real parts, then $G(z)$ is analytic at $z = u_1$. Since $p_0(z)$ is not identically zero, this implies that $(z - u_1)^{\alpha_1}$ is meromorphic at $z = u_1$, hence α_1 is an integer, a contradiction. The converse implication is trivial. ■

Another sequence that can be treated in a similar fashion is $n^{\alpha n}$. So far, its non-holonomicity was only known for rational α [6].

Proposition 4.2. *For every non-zero complex α , the sequence $(n^{\alpha n})_{n \geq 1}$ is not holonomic.*

Proof. Let $F(z)$ be as usual. It is easily verified that $(z+1)^{-\alpha(z+1)}F(z+1)$ satisfies the growth conditions of Theorem 2.2. Appealing to condition (A1) completes the proof. ■

Now let us consider rational functions in n and e^n with real coefficients. By Khovanskii's theorem and some commutative algebra, we can precisely characterize the holonomic sequences of this kind, even if they contain several exponentials.

Theorem 4.3. *Let $\alpha_1, \dots, \alpha_m$ be positive real numbers such that $\log \alpha_1, \dots, \log \alpha_m$ are linearly independent over \mathbb{Q} . If $f \in \mathbb{R}(x, \alpha_1^x, \dots, \alpha_m^x)$ is such that the sequence $(f(n))_{n \geq 0}$ is holonomic, then $f(x) \in \mathbb{R}(x)[\alpha_1^x, \alpha_1^{-x}, \dots, \alpha_m^x, \alpha_m^{-x}]$.*

Proof. We write $\mathbf{y} = (y_1, \dots, y_m)$ for a vector of indeterminates and take some element $g(x, \mathbf{y}) = A(x, \mathbf{y})/B(x, \mathbf{y})$ of $\mathbb{R}(x, \mathbf{y})$ with the property that there is n_0 such that $(g(n, \alpha_1^n, \dots, \alpha_m^n))_{n \geq n_0}$ is a holonomic sequence. Without loss of generality the polynomials A and B are coprime, and B depends on y_m . By the closure of holonomic sequences under multiplication, we may also assume that B is irreducible. Moreover, $B \notin \mathbb{R}[x] \cup \mathbb{R}[x]y_m$, or else there is nothing to show. Since this implies that Ay_m and B are coprime too, there are polynomials $L(x, \mathbf{y})$ and $M(x, \mathbf{y})$ such that

$$L(x, \mathbf{y})A(x, \mathbf{y})y_m + M(x, \mathbf{y})B(x, \mathbf{y}) = R(x, y_1, \dots, y_{m-1}). \quad (4.2)$$

Here, R is the resultant of the two polynomials Ay_m and B , viewed as polynomials in the variable y_m . Dividing (4.2) by B , we may thus assume that the numerator of g does not depend on y_m . Pulling out the coefficient of the highest power of y_m , we may at last assume that our g is of the form

$$g(x, \mathbf{y}) = \frac{R(x, y_1, \dots, y_{m-1})}{h_0 + \dots + h_{D-1}y_m^{D-1} + y_m^D},$$

where $D > 0$, $h_j \in \mathbb{R}(x, y_1, \dots, y_{m-1})$, $h_0 \not\equiv 0$, and the denominator is irreducible. Suppose now that $g(n, \alpha_1^n, \dots, \alpha_m^n)$ satisfies (1.1) for large n . By Corollary 2.5 (in fact, its obvious extension to $[n_0, \infty]$), it follows that

$$0 = \sum_{k=0}^d p_k(x)g\left(x+k, \alpha_1^{x+k}, \dots, \alpha_m^{x+k}\right) = \sum_{k=0}^d p_k(x)g_k(x, \alpha_1^x, \dots, \alpha_m^x),$$

for large real x . Here, we have defined

$$g_k(x, \mathbf{y}) := g\left(x+k, \alpha_1^k y_1, \dots, \alpha_m^k y_m\right).$$

The multiplicative independence of the numbers $\alpha_1, \dots, \alpha_m$ implies the algebraic independence of the functions $\alpha_1^x, \dots, \alpha_m^x$. (Indeed, substituting these functions in a non-zero polynomial P , we get a sum $P(x, \alpha_1^x, \dots, \alpha_m^x)$ of monomials of the form $p(x)\beta^x$, where $\beta = \alpha_1^{n_1} \cdots \alpha_m^{n_m}$, with exactly one monomial having maximal β . This monomial cannot be cancelled.) It follows that

$$\sum_{k=0}^d p_k(x)g_k(x, \mathbf{y}) = 0.$$

Since the numerators of the g_k do not depend on y_m , there must be $k > 0$ such that the denominators of g_0 and g_k , as polynomials in $\mathbb{R}(x, y_1, \dots, y_{m-1})[y_m]$, have a common root in the algebraic closure of $\mathbb{R}(x, y_1, \dots, y_{m-1})$. (Essentially, we are once more appealing to condition (A1).) But as these denominators are irreducible, the only way that they can share a root is if they coincide, up to a multiplicative “constant” from $\mathbb{R}(x, y_1, \dots, y_{m-1})$. The leading coefficients with respect to y_m of the denominators of g_0 and g_k are 1 and α_m^{kD} , respectively, hence identifying the coefficients of y_m^0 yields

$$h_0(x, y_1, \dots, y_{m-1}) = \alpha_m^{-kD} h_0\left(x, \alpha_1^k y_1, \dots, \alpha_{m-1}^k y_{m-1}\right).$$

But this equation cannot hold, since comparing the coefficient of any monomial (with non-zero coefficient) gives rise to a multiplicative dependence relation between the α_i . ■

Singer and van der Put [24] have shown that the reciprocal $(1/a_n)$ of a holonomic sequence (a_n) is not holonomic unless (a_n) is a periodic interlacement of hypergeometric sequences. Thus, in the special case $m = 1$ we can close the above proof after showing that the numerator of $g(x, y_1)$ is without loss of generality independent of α_1^x .

To see that the linear independence assumption in Theorem 4.3 is necessary, suppose that

$$c_1 \log \alpha_1 + \cdots + c_m \log \alpha_m = 0,$$

where c_1, \dots, c_m are integers, and the α_i are without loss of generality ordered such that

$$c_1, \dots, c_k \geq 0 \quad \text{and} \quad c_{k+1}, \dots, c_m \leq 0.$$

Then the sequence $g((n, \alpha_1^n, \dots, \alpha_m^n))_{n \geq 1}$ corresponding to

$$g(x, \mathbf{y}) = \frac{1}{1 + y_1^{c_1} \cdots y_k^{c_k} - y_{k+1}^{-c_{k+1}} \cdots y_m^{-c_m}}$$

is the constant sequence $(1)_{n \geq 1}$.

5. Miscellaneous Results: Derivatives, Algebraic Functions, and Interlacing

In this section we collect several results, mostly consequences of Khovanskii's theorem.

Take some functions $f(x)$ and $g(x)$ defined for $x > 0$ such that $(f(n))_{n \geq 1}$ and $(g(n))_{n \geq 1}$ are holonomic sequences. Which operations preserve holonomicity? Clearly, the sequences arising from $f(x) + g(x)$ and $f(x)g(x)$ are holonomic; we have been using this fact all the time. Since we assume that $f(x)$ is defined for all positive real numbers, there are some other natural operations that would not make sense for sequences (a_n) in general. Integration does not necessarily lead to a function defining a holonomic sequence, as shown by $f(x) = 1/x$. Concerning differentiation, we note the following result.

Proposition 5.1. *Let $f(x)$ be an elementary function that is a quotient of D-finite functions defined for $x > 0$. If the sequence $(f(n))_{n \geq 1}$ is holonomic, then the sequence $(f'(n))_{n \geq 1}$ is holonomic too.*

Proof. Assume that the sequence $a_n = f(n)$ satisfies (1.1). By Corollary 2.5, this implies

$$f(x+d) = -p_d(x)^{-1} \sum_{k=0}^{d-1} p_k(x)f(x+k), \quad x \geq 1.$$

From this relation and Leibniz's rule it follows easily that, for every $l \geq d$, the function $f(x+l)$ is in the vector space spanned by $f(x), \dots, f(x+d-1)$ and $f'(x), \dots, f'(x+d-1)$ over $\mathbb{R}(x)$. Since this vector space has dimension at most $2d$, there must be a linear relation among $f'(x), \dots, f'(x+2d)$. ■

Another natural question, which goes slightly beyond closed-form sequences, is whether sequences that satisfy some algebraic equation can be holonomic. It is clear that periodic interlacings of rational functions are holonomic, but no other algebraic holonomic sequences are known. For instance, we know from Section 3 that the sequences n^α with $\alpha \in \mathbb{Q} \setminus \mathbb{Z}$ are not-holonomic, a fact already established by Gerhold [6] with a number-theoretic argument.

Theorem 5.2. *Let $f(z)$ be an algebraic function that is analytic in a neighborhood of $[1, \infty[$. Then the sequence $(f(n))_{n \geq 1}$ is holonomic if and only if $f(z)$ is a rational function.*

Proof. Let $F(z) = F(z; f, p_0, \dots, p_d)$ be as usual. Let \mathcal{V} be the algebraic variety generated by $\{(z, F(z)): z \geq 1\}$. By Bézout's theorem, the points $(z, F(z))$ stay in one irreducible component of \mathcal{V} for large real z (otherwise two components would have infinitely many intersections). As we suppose that $F(z)$ vanishes for $z \in \mathbb{N}$, this component has infinitely many intersections with the real axis, hence it equals the real axis, again by Bézout's theorem. But this shows that $F(z)$ vanishes identically. Now observe that $f(z)$ has an analytic continuation to a slit plane with finitely many singularities and branch points. If all the singularities are poles, then $f(z)$ is rational [8, p. 218], and we are done. Otherwise, we choose a singularity z_0 , not a pole, such that $z_0 + 1, z_0 + 2, \dots$ are either poles or not singularities of $f(z)$. Then we can appeal to condition (A1), since $F(z)$ is not meromorphic at $z = z_0$. ■

A weaker assumption would be that the points (n, a_n) satisfy some algebraic equation. The result should then be that a holonomic sequence arises only for a periodic interlacement of rational functions. The problem with the above proof is that the points (n, a_n) may lie in several components of the variety generated by $\{(n, a_n): n \in \mathbb{N}\}$. We can show as above that each of these components gives rise to a rational function, but what is missing is an argument that these rational functions must be interlaced in a periodic way. The special case where the rational functions are constants is doable, though:

Proposition 5.3. *If a holonomic sequence has only finitely many distinct values, then it is eventually periodic.*

Proof. It is known [16, Exercise VIII.158] that this holds if the sequence satisfies a linear recurrence with *constant* coefficients. We show that the result for a sequence $(a_n)_{n \geq 1}$ satisfying a recurrence of the form (1.1) follows from this special case. Let N be the maximum of the degrees of the p_k 's, and denote the coefficient of n^j in $p_k(n)$ by b_{kj} . Then

$$0 = \sum_{k=0}^d p_k(n)a_{n+k} = \sum_{j=0}^N L_j(n)n^j$$

for $n \geq 1$, where

$$L_j(n) := \sum_{k=0}^d b_{kj}a_{n+k}.$$

By assumption, the set of $N + 1$ dimensional vectors

$$\mathcal{L} = \{(L_0(n), \dots, L_N(n)): n \geq 1\}$$

is finite. Since a non-zero polynomial has only finitely many roots, each non-zero vector from \mathcal{L} can occur only for finitely many n . In particular, $L_N(n)$ must vanish for large n . By definition of N , some b_{jN} is non-zero, which yields a non-trivial constant coefficient recurrence for (a_n) . ■

Szegő has proved that a power series with finitely many distinct coefficients is either a rational function or has the unit circle as a natural boundary [17]. This provides another way to achieve the reduction to constant coefficients in the preceding proof.

Finally, we briefly address the question whether non-holonomic sequences can be put together to produce a holonomic sequence.

Proposition 5.4. *Suppose that the sequence a_n is a (not necessarily periodic) interlacement of the sequences $\log n$ and \sqrt{n} . Then a_n is non-holonomic.*

Proof. Suppose that a_n satisfies a recurrence of order d . Let us color the indices n where $\log n$ occurs black and the others white. In this coloring of \mathbb{N} , there must be some pattern of length $d+1$ that occurs infinitely often. Hence one of the 2^{d+1} functions

$$F(x) = p_0(x) \left\{ \begin{array}{l} \sqrt{x} \\ \log x \end{array} \right\} + \cdots + p_d(x) \left\{ \begin{array}{l} \sqrt{x+d} \\ \log(x+d) \end{array} \right\}$$

has infinitely many zeros. By Corollary 2.3 (or Corollary 2.5), it vanishes identically, but this can once again be refuted by condition (A1). Hence a_n is not holonomic. ■

6. The Riemann Zeta Function

In this section we consider (subsequences of) the sequence $(\zeta(n))_{n \geq 2}$, where $\zeta(s) = \sum_{l \geq 1} l^{-s}$ denotes the Riemann zeta function. The right-hand side of Euler's formula [14]

$$\sum_{k \geq 1} \zeta(2k) z^{2k} = \frac{1}{2} - \frac{\pi z}{2} \cot \pi z$$

reveals infinitely many singularities, hence $(\zeta(2n))_{n \geq 1}$, and hence $(\zeta(n))_{n \geq 2}$, are not holonomic. The values of the zeta function at *odd* integers have attracted the attention of number theorists for (at least) the last decades, the most well-known result being Apéry's proof [1, 23] of the irrationality of $\zeta(3)$. Molteni [14] has noted that

$$\sum_{k \geq 1} \zeta(2k+1) z^{2k} = z^{-1} \sum_{k \geq 1} \zeta(2k) z^{2k} - \psi(z+1) - \gamma, \quad (6.1)$$

where $\psi = \Gamma'/\Gamma$ is the digamma function, and γ is Euler's constant. Since $\psi(z)$ is analytic in the right-half plane, this shows that $(\zeta(2n+1))_{n \geq 1}$ is not holonomic, again by the infinitely many singularities of its generation function. Molteni [14] has already observed that (6.1) implies that $(\zeta(2n+1))_{n \geq 1}$ satisfies no linear recurrence with constant coefficients[‡].

We will now extend these two results to sequences $(\zeta(b_n))_{n \geq 1}$, where $(b_n)_{n \geq 1}$ is an increasing integer sequence with bounded gap lengths. None of our conditions for property (A) is applicable. In particular, the zeta function is meromorphic in \mathbb{C} , violating condition (A1). This prompts us to look for another argument. Considering the defining series of ζ summand by summand, we will deduce a full rank set of linear relations for the $p_k(n)$. As an analogy with the theory of transcendental numbers, we mention that the proof is evocative of the well-known fact that a rapidly converging series can be very useful for establishing the transcendence of a number.

[‡] We thank Tanguy Rivoal for bringing Molteni's paper to our attention.

Theorem 6.1. *Let $2 \leq b_1 < b_2 < b_3 < \dots$ be a sequence of integers such that the gap lengths $b_{n+1} - b_n$ are bounded. Then the sequence $(\zeta(b_n))_{n \geq 1}$ is not holonomic.*

Proof. Suppose that there are polynomials $p_k(x)$ with real coefficients such that

$$\sum_{k=0}^d p_k(n) \zeta(b_{n+k}) = 0$$

holds for all integers $n \geq 1$. We show that every $p_k(x)$ is the zero polynomial. By the estimate

$$\sum_{l \geq L} l^{-s} \leq L^{-s} + \int_L^\infty t^{-s} dt = O(L^{-s}), \quad s \rightarrow \infty,$$

we find

$$\sum_{k=0}^d p_k(n) = - \sum_{k=0}^d p_k(n) \sum_{l \geq 2} \frac{1}{l^{b_{n+k}}} \rightarrow 0, \quad n \rightarrow \infty,$$

which implies that $\sum_{k=0}^d p_k(x)$ is zero. We iterate this argument as follows: From

$$\sum_{k=0}^d p_k(n) \left(1 + \frac{1}{2^{b_{n+k}}} \right) = \sum_{k=0}^d p_k(n) \frac{1}{2^{b_{n+k}}} = - \sum_{k=0}^d p_k(n) \sum_{l \geq 3} \frac{1}{l^{b_{n+k}}},$$

we obtain

$$\sum_{k=0}^d p_k(n) \frac{1}{2^{b_{n+k}-b_n}} = -2^{b_n} \sum_{k=0}^d p_k(n) \sum_{l \geq 3} \frac{1}{l^{b_{n+k}}} \rightarrow 0, \quad n \rightarrow \infty.$$

Now for every $n \geq 1$ and $0 \leq k \leq d$ we have $b_{n+k} - b_n \in \{0, 1, 2, \dots, r\}$ for an absolute constant $r \in \mathbb{N}$, because the gap lengths are bounded. Since there are only finitely many $(\leq (r+1)^d)$ possibilities for the $(d+1)$ -tuple $(b_{n+k} - b_n : 0 \leq k \leq d)$, there exists an n_0 such that the numbers $n > n_0$ can be partitioned into finitely many infinite sequences S_1, S_2, \dots, S_p so that $(b_{n+k} - b_n : 0 \leq k \leq d)$ is constant for n running through S_i . Thus, we can conclude as before that

$$\sum_{k=0}^d p_k(n) \frac{1}{2^{b_{n+k}-b_n}} = 0, \quad n \in S_i.$$

This is true for every S_i , and, dividing by 2^{b_n} , we get that

$$\sum_{k=0}^d p_k(n) \frac{1}{2^{b_{n+k}}} = 0, \quad n > n_0.$$

Using this, we can prove in the same way that

$$\sum_{k=0}^d p_k(n) \frac{1}{3^{b_{n+k}}} = 0, \quad n > n_1,$$

with n_1 sufficiently large, and so on. We see that

$$\sum_{k=0}^d p_k(n) \frac{1}{L^{b_{n+k}}} = 0, \quad n \text{ large},$$

for every $L = 1, 2, \dots, d+1$. Using that the generalized Vandermonde matrix $(L^{-b_{n+k}})$, $1 \leq L \leq d+1$ and $0 \leq k \leq d$, is invertible [5] (we have no repetition in the sequence $(b_n)_{n \geq 1}$), we obtain

$$p_k(n) = 0, \quad n \text{ large},$$

for $k = 0, 1, \dots, d$. Thus every $p_k(x)$ is identically zero. \blacksquare

7. Comments and Remarks

When trying to establish property (A), we have to show that a certain function F does not vanish identically. A possible limit to this endeavor is hinted at by the celebrated results of Richardson [18] and Caviness [2]. Recall that they have shown that it is an algorithmically undecidable problem to determine whether a given function from certain classes of elementary functions vanishes identically. On the one hand, their functions have the absolute value function as a building block, which we do not consider; on the other hand, they assume a closed form expression, while our F contains arbitrary polynomials.

As for property (B), we remark that there is an extension of Carlson's classical result due to Malliavin and Rubel [13, 21].

Applying Khovanskii's theorem only to quotients of D -finite functions severely hampers the power of the theorem. Maybe the order of zeros of elementary functions is bounded, which would make this restriction superfluous. An observation that might prove useful is that all elementary functions are differentially algebraic, i.e., $P(z, f, f', \dots, f^{(k)}) = 0$ for a non-zero polynomial P (in $k+2$ variables). Hopefully this representation can be used to get an upper bound on the order of a zero of f . However, examples such as $(f')^2 + f \cdot f'' = 0$ show that for algebraic differential equations the boundedness cannot be established by the same straightforward argument as for D -finite functions. Rubel [19, Problem 20] [20, Problem 29] asks which sequences (z_1, z_2, \dots) (counting multiplicities) can serve as zero sets of entire differentially algebraic functions and remarks that the case $(1, 2, 3, \dots)$ is still undecided.

Furthermore, we remark that, surprisingly, in the case of integer sequences our restriction to sequences defined via elementary functions is not that restrictive after all. Laczkovich and Ruzsa [11] proved that every integer sequence can be represented in the form $(f(n))_{n \geq 1}$, where f is a so-called naive-elementary function. However, this representation relies on the infinitude of zeros of the sine function, conflicting with the requirements of Khovanskii's theorem.

We conclude our final remarks with four open problems.

- (i) Besides those described in Theorem 4.3, which other elementary (in Khovanskii's sense) functions define holonomic sequences? (Note that $(\sin \alpha n)_{n \geq 0}$ clearly is holonomic.)

- (ii) Is a sequence satisfying an algebraic equation holonomic if and only if it is a periodic interlacement of rational functions?
- (iii) Can our method be pushed forward to accomodate recurrences of more general shapes and multivariate sequences?
- (iv) Let $f(x)$ be a non-rational elementary function such that $(f(n))_{n \geq 1}$ is holonomic. If $s(x)$ is some function such that $(f(s(n)))_{n \geq 1}$ is also holonomic, does it follow that $s(x)$ is linear? (Of course, an appropriate condition to exclude trivial cases like $s = f^{-1}$ is needed.)

References

1. R. Apéry, Irrationalité de $\zeta(2)$ et $\zeta(3)$, Astérisque **61** (1979) 11–13.
2. B.F. Caviness, On canonical forms and simplification, J. ACM **17** (1970) 385–396.
3. P. Flajolet, S. Gerhold, and B. Salvy, On the non-holonomic character of logarithms, powers and the n th prime function, Electron. J. Combin. **11** (2005) 1–16.
4. P. Flajolet, S. Gerhold, and B. Salvy, Asymptotic analysis of the generating functions of certain closed form sequences, in preparation.
5. F.R. Gantmacher, The Theory of Matrices II, Chelsea Publishing Co., New York, 1959.
6. S. Gerhold, On some non-holonomic sequences, Electron. J. Combin. **11** (2004) 1–8.
7. S. Gerhold, Combinatorial sequences: non-holonomicity and inequalities, PhD thesis, Johannes Kepler University Linz, 2005.
8. P. Henrici, Applied and Computational Complex Analysis, Vol. 1, John Wiley, New York, 1974.
9. A.G. Khovanskii, Fewnomials, AMS Press, Providence, RI, 1991.
10. M. Klazar, Non-holonomicity of the sequence $\log 1, \log 2, \log 3, \dots$, available at: arXiv: math.CO/0502141.
11. M. Laczkovich and I. Ruzsa, Elementary and integral-elementary functions, Illinois J. Math. **44** (2000) 161–182.
12. L. Lipshitz, D -finite power series, J. Algebra **122** (1989) 353–373.
13. P. Malliavin and L. Rubel, On small entire functions of exponential type with given zeros, Bull. Soc. Math. France **89** (1961) 175–206.
14. G. Molteni, Some arithmetical properties of the generating power series for the sequence $\{\zeta(2k+1)\}_{k=1}^{\infty}$, Acta Math. Hungar. **90** (1-2) (2001) 133–140.
15. G. Pólya and G. Szegő, Problems and Theorems in Analysis I, Classics in Mathematics, Springer-Verlag, Berlin, 1998.
16. G. Pólya and G. Szegő, Problems and Theorems in Analysis II, Classics in Mathematics, Springer-Verlag, Berlin, 1998.
17. R. Remmert, Classical Topics in Complex Function Theory, Springer, 1998.
18. D. Richardson, Some undecidable problems involving elementary functions of a real variable, J. Symbolic Logic **33** (4) (1968) 514–520.
19. L.A. Rubel, Some research problems about algebraic differential equations, Trans. Amer. Math. Soc. **280** (1983) 43–52.
20. L.A. Rubel, Some research problems about algebraic differential equations II, Illinois J. Math. **36** (1992) 659–680.
21. L.A. Rubel, Entire and Meromorphic Functions, Springer-Verlag, New York, 1996.

22. R.P. Stanley, *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge, 1998.
23. A. van der Poorten, A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$, *Math. Intelligencer* **1** (1979) 195–203.
24. M. van der Put and M.F. Singer, Galois theory of difference equations, *Lecture Notes in Mathematics*, Vol. 1666, Springer-Verlag, Berlin, 1997.