

FIRST-ORDER FACTORS OF LINEAR MAHLER OPERATORS

FRÉDÉRIC CHYZAK, THOMAS DREYFUS, PHILIPPE DUMAS,
AND MARC MEZZAROBBA

Dedicated to the memory of Marko Petkovšek.

ABSTRACT. We develop and compare two algorithms for computing first-order right-hand factors in the ring of linear Mahler operators $\ell_r M^r + \dots + \ell_1 M + \ell_0$ where ℓ_0, \dots, ℓ_r are polynomials in x and $Mx = x^b M$ for some integer $b \geq 2$. In other words, we give algorithms for finding all formal infinite product solutions of linear functional equations $\ell_r(x)f(x^{b^r}) + \dots + \ell_1(x)f(x^b) + \ell_0(x)f(x) = 0$.

The first of our algorithms is adapted from Petkovšek's classical algorithm for the analogous problem in the case of linear recurrences. The second one proceeds by computing a basis of generalized power series solutions of the functional equation and by using Hermite–Padé approximants to detect those linear combinations of the solutions that correspond to first-order factors.

We present implementations of both algorithms and discuss their use in combination with criteria from the literature to prove the differential transcendence of power series solutions of Mahler equations.

CONTENTS

1. Introduction	2
Context and motivation	2
Problem	3
Contribution	3
Related work	4
Outline	5
Notation	6
2. Structure of the solution set of the Riccati equation	6
2.1. Basic notions of difference algebra	6
2.2. First-order factors and their solutions	7
2.3. Parametrization of the solution set	8
3. Generalized series solutions of the linear equation	10
3.1. The difference ring \mathfrak{D}	10
3.2. Bounds	12
3.3. Puiseux-hypergeometric solutions	14
4. Ramified rational solutions to the Riccati equation	15

© This work is licensed under a Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

For the purpose of Open Access, a CC-BY public copyright licence has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission.

Supported in part by the French ANR grant *De rerum natura* (ANR-19-CE40-0018) and by the French-Austrian ANR-FWF grant EAGLES (ANR-22-CE91-0007 & FWF-I-6130-N). The IMB receives support from the EIPHI Graduate School (contract ANR-17-EURE-0002).

4.1. Hypergeometric elements	15
4.2. Similarity classes	16
5. Degree bounds for rational solutions	18
6. Mahlerian variant of Petkovšek’s algorithm	20
6.1. Petkovšek’s classical algorithm	20
6.2. Roques’s algorithm for order 2	21
6.3. A new algorithm for higher-order Mahler equations	22
6.4. Existence and computation of bounded Gosper–Petkovšek forms	27
6.5. Optimizations	28
7. Alternative algorithm by Hermite–Padé approximation	33
7.1. Approximate syzygies	34
7.2. From unstructured to structured syzygies	36
7.3. Ramified rational solutions	42
7.4. Algorithm by sieving candidates	42
7.5. A supplementary remark on the rank of relations module	48
7.6. Rational solving of the linear Mahler equation	49
8. Implementation and benchmark	50
8.1. Implementation	50
8.2. Examples	51
8.3. Discussion of the timings	55
9. Differential transcendence of Mahler functions	56
References	60

1. INTRODUCTION

Context and motivation. Mahler equations are a type of functional equations that, for some fixed integer $b \geq 2$, relate the iterates $y(x)$, $y(x^b)$, $y(x^{b^2})$, \dots , of an unknown function y under the substitution of x^b for x . Mahler originally considered nonlinear multivariate equations of this type in his work in transcendence theory in the 1920s. The focus later shifted to linear univariate equations, in relation to the study of automatic sequences. More recently, questions in difference Galois theory related to the existence of nonlinear differential equations satisfied by solutions of difference equations of various types led to a revival of the topic.

The present article continues a line of work initiated in our earlier publication (CDDM 2018), to which we refer for more context. This work started when, back in December 2015, the second author asked about making effective a differential transcendence criterion introduced in (Dreyfus, Hardouin, and Roques 2018). This criterion boils down to determining the rational function solutions of a nonlinear Mahler equation analogous to the Riccati equation associated to a linear differential equation. Equivalently, this can be viewed as computing the first-order right-hand factors with rational function coefficients of the linear Mahler operator underlying the Mahler equation. We come back to this original motivation in §9 and postpone to that section a more detailed discussion of differential transcendence and differential algebraic independence criteria based on Mahler equations.

A second motivation for studying first-order factors is that, in the differential case, Beke’s method (Markoff 1891; Bendixson 1892; Beke 1894)¹ reduces the problem

¹See (Bostan, Rivoal, and Salvy 2023) for more on the history of this method.

of factoring a differential equation into irreducible factors to that of finding the first-order right-hand factors of some auxiliary equations. One may expect that the method adapts to the Mahler case, so the present work paves the way to future factorization algorithms.

Problem. We now fix the notation that we will use throughout. For some fixed integer $b \geq 2$, we are interested in equations involving the *Mahler operator* M with regard to b , which is defined as mapping any function $y(x)$ to $y(x^b)$. For $r \in \mathbb{N}$ and polynomials $\ell_i \in \mathbb{K}[x]$ with coefficients in a field \mathbb{K} to be fixed in a moment, we consider more specifically the equations

$$(L) \quad \ell_r M^r y + \cdots + \ell_1 M y + \ell_0 y = 0,$$

$$(R) \quad \ell_r u \cdots M^{r-1} u + \cdots + \ell_2 u M u + \ell_1 u + \ell_0 = 0,$$

which we will refer to as, respectively, the *linear Mahler equation* (L) and the *Riccati Mahler equation* (R). We also attach to the linear equation the corresponding linear operator

$$L = \ell_r M^r + \cdots + \ell_1 M + \ell_0.$$

Throughout, we further assume that both polynomials ℓ_0 and ℓ_r are nonzero and that L is primitive, that is, that the family of the ℓ_i has gcd 1. We write d for the maximal degree $\max_k \deg \ell_k$ of the coefficients. As usual, we call r the *order* of the linear Mahler equation and of the underlying operator L .

Our goal is to find the right-hand factors $M - u$ of L where u is a ramified rational function. Equivalently we want to compute the *hypergeometric* solutions of L , that is, the solutions y of L that satisfy a first-order linear Mahler equation with ramified rational coefficients. Informally, the link between both equations is the relation $u = My/y$. However, general solutions of (L) may live in a ring containing zero divisors, so that quotients are not always well defined (see the construction of \mathfrak{D} in §3.1).

Developing algorithms for solving the Riccati Mahler equation requires solving linear Mahler equations in various domains. To this end, we build on results from (CDDM 2018). There, polynomials, rational functions, and series were defined with coefficients in a computable subfield \mathbb{K} of \mathbb{C} , so we continue with this assumption here.

Contribution. This article provides algorithms to compute the rational solutions of Riccati Mahler equations. We develop two approaches.

The first one is adapted from the classical algorithm by Petkovšek (1992) for finding the hypergeometric solutions of a linear recurrence equation. Petkovšek's algorithm searches for first-order factors of difference operators in a special form called the Gosper–Petkovšek form. The existence of Gosper–Petkovšek forms relies on the fact that for any two nonzero polynomials A and B , the set of integers i such that $A(x)$ and $B(x+i)$ have a nontrivial common divisor is finite. As this is not true when the shift is replaced by the Mahler operator, we need to slightly depart from the classical definition of Gosper–Petkovšek forms. Doing so, we obtain a first complete algorithm for finding first-order factors of Mahler equations of arbitrary order. Note that Roques (2018) recently gave a slightly different adaptation of Petkovšek's algorithm to the case of Mahler equations of order two. Like Petkovšek's, these algorithms have to consider an exponential number of separate subproblems

in the worst case. We discuss several ways of pruning the search space to mitigate this issue in practice.

Our second algorithm avoids the combinatorial search phase entirely, at the price of a worst-case exponential behavior of a different nature. It is based on a relaxation of the problem that can be solved using Hermite–Padé approximants². The idea is to search for series solutions y of the linear Mahler equation that make My/y rational. Roughly speaking, we first compute a basis (y_1, \dots, y_t) of series solutions, then search for linear combinations with polynomial coefficients of $y_1, \dots, y_t, My_1, \dots, My_t$ that vanish to a high approximation order σ , and finally isolate, among these relations, those corresponding to hypergeometric solutions by solving a polynomial system. Though we are not aware of any exact analogue of our algorithm in the literature, variants of the same basic idea have been used by several authors in the differential case (see below for references).

In order to state the algorithms and justify their correctness, it is useful to work in a ring containing “all” solutions, or at least all solutions needed in the discussion, of the linear equation (L). Instead of appealing for this to the general Picard–Vessiot theory of linear difference equations, we introduce a notion of *1-universal extension* that suffices for our purposes and construct a 1-universal ring whose elements (unlike those of Picard–Vessiot rings of Mahler equations) admit simple representations as formal series. In passing we define a suitable notion of Mahler-hypergeometric function and establish its basic properties.

We compare the performance of our two approaches based on an implementation³ in Maple and observe that the second algorithm turns out to be more efficient in practice on examples from the mathematical literature. Finally, we use this implementation in combination with criteria such as the one mentioned earlier to prove several differential algebraic independence results between series studied in the literature.

Related work. To the best of our knowledge, the problem we consider here was first discussed in the doctoral dissertation of Dumas (1993, §3.6), which contains an incomplete method for finding hypergeometric solutions of Mahler equations. Dumas’ method is somewhat reminiscent of Petkovšek’s algorithm, but lacks a proper notion of Gosper–Petkovšek form. Roques (2018, §6), as already mentioned, describes a complete analogue of Petkovšek’s algorithm for Mahler equations, but restricts himself to equations of order two. We are not aware of any other reference dealing specifically with the factorization of Mahler operators.

However, it is natural to try and adapt to Mahler equations algorithms that apply to differential equations or to difference equations in terms of the shift operator. In the differential case, factoring algorithms are a classical topic, going back at least to Fabry’s time (1885, §V); we refer to (Bostan, Rivoal, and Salvy 2023, §1.4) for a well-documented overview. The second of our algorithms, using Hermite–Padé approximation, is related to methods known from the differential case. Most similar to our approach is maybe an unpublished article by Bronstein and Mulders (n.d. [1999?], §4) where they present a heuristic method for solving Riccati differential equations based on Padé (not Hermite–Padé) approximants of series solutions with

²We find some pleasant irony in our application of Hermite–Padé approximants to problems on Mahler operators, after Mahler himself has introduced similar approximants in his work (Mahler 1968).

³This package by Ph. Dumas will be made available in the future.

indeterminate coefficients. The same idea appears in works by Pflügel (1997, §2.5.2) and by van der Put and Singer (2003, §4.1), though neither of these references discusses in detail how to deal with drops in the degree of candidate solutions for special values of the parameters. The core idea of enforcing the vanishing of high-order terms of series solutions of the Riccati equation so as to reduce to the solution of polynomial equations in a number of unknowns bounded by the order of the differential equation already appears in Fabry’s thesis.

In the shift case, the classical algorithm for finding hypergeometric solutions is that of Petkovšek (1992). It is the direct inspiration for our first method. Petkovšek’s algorithm is itself based on Gosper’s hypergeometric summation algorithm (Gosper 1978), and was previously adapted to q -difference equations in (Abramov and Petkovšek 1995; Abramov, Paule, and Petkovšek 1998). Van Hoeij and Cluzeau (van Hoeij 1999; Cluzeau and Van Hoeij 2004) later proposed alternative algorithms that are faster in practice; it seems likely that the idea would also be relevant in the Mahler case, but we do not explore this question here.

Another line of work aims at unified algorithms for linear functional operators of various types based on the formalism of Ore polynomials (e.g., Bronstein and Petkovšek 1993). Factoring algorithms, however, remain specific to each individual type of equation. In addition, even methods that do apply to almost all types of Ore operators sometimes fail in the Mahler case because the commutation $Mx = x^bM$ does not preserve degrees with respect to x .

Turning now to the structure and computation of generalized series solutions of linear Mahler equations, Roques’s discussion of the local exponents of Mahler systems (2018, §4) forms the basis for our §3. Further developments (not used here) include recent work by Roques (2023), Faverjon and Roques (2022), and Faverjon and Poulet (2022).

Outline. The “structural” results about the solution spaces of Mahler equations come first in the text. We first generalize our Mahlerian problem to the context of difference rings, including the case of nonsurjective morphisms: in §2, we derive a parametric description of the right-hand factors of a difference operator (Theorem 2.8). To accommodate the Mahler operator in the previous generalized framework, we then introduce in §3 in an explicit difference ring \mathfrak{D} containing all the solutions of the linear Mahler equations that are needed in the parametric descriptions of the solutions u to Riccati Mahler equations. In that section, we also define classes of F -hypergeometric solutions for various difference fields F and, for F the field of Puiseux series, we describe the solution set of the Riccati equation in the Puiseux series as a partition according to the coefficient of their valuation term (Theorem 3.14). To describe the ramified rational solutions of the Riccati equation, we then change F to the field of ramified solutions in §4, where we obtain a partition (Theorem 4.5) that is finer than the partition induced by the partition in the Puiseux series on ramified solutions. In preparation for the algorithms, the technical §5 presents bounds on the degree of numerators and denominators of rational solutions to a Riccati Mahler equation (Proposition 5.2).

We continue with two algorithmic approaches. We review Petkovšek’s classical algorithm for the shift case and Roques’s analogue for order 2 in the Mahler case before developing our generalization in §6. We first propose a brute-force algorithm (Algorithm 1), which we prove to be correct (Theorem 6.3), and then propose several pruning criteria to improve its exponential behavior, leading to an improved

algorithm (Algorithm 3). We then develop our relaxation method based on Hermite–Padé approximations in §7: our algorithm (Algorithm 4) produces successive sets of parametrized candidates, each containing all true solutions, until it stabilizes on true solutions only; justifying termination and correctness relies on an accurate description of the limit (Lemma 7.2) and on a delicate primary decomposition argument (Theorem 7.15).

To conclude, we present examples of applications and timings in §8, where we can observe that our relaxation method beats the other by combinatorial exploration in a number of natural examples. We finally apply our implementation in order to prove properties of differential transcendence on examples in §9.

Notation. Henceforth, we will denote the field of Puiseux series over a field \mathbb{L} by $\mathbb{L}((x^{1/*})) = \bigcup_{q \geq 1} \mathbb{L}((x^{1/q}))$. Similarly, the field $\bigcup_{q \geq 1} \mathbb{L}((x^{1/q}))$ of ramified rational functions will be denoted $\mathbb{L}(x^{1/*})$.

Occasionally, we use the notation $\mathbb{L}[x]\langle M \rangle$ to denote the algebra generated by M over $\mathbb{L}[x]$ and satisfying the relation $Mf(x) = f(x^b)M$ for any $f \in \mathbb{L}[x]$. Similar definitions apply for other coefficient domains.

All of the operators and functions M , val , \ln , \log take precedence over additions and products, which means $My/y = (My)/y$, $\text{val } ab = (\text{val } a)b$, etc. We write Ly for the application of an operator L to a function, or $L(x, M)y(x)$ if needed.

We write $S_{\neq 0}$ for a given set S containing 0 to denote $S \setminus \{0\}$. The set S will be the set \mathbb{N} of natural numbers, a field, a vector space, a cone, a Cartesian product, etc.

2. STRUCTURE OF THE SOLUTION SET OF THE RICCATI EQUATION

The notions introduced in this section will be used when solving Riccati Mahler equations. As they do not depend on the specific choice of the Mahler operator, we write them in the broader generality of difference rings.

2.1. Basic notions of difference algebra. A difference ring is commonly defined as a pair (D, σ) formed by a commutative ring with identity D and an automorphism σ of D ; see, e.g., (Cohn 1965) or (van der Put and Singer 1997), and a difference field is a difference ring that is a field. An example is the field $\mathbb{K}((x^{1/*}))$ of Puiseux series, equipped with the Mahler operator M . However, we relax the definition to accept a map σ that is only an injective endomorphism (*cf.* Wibmer 2013), since we intend to consider as well the restriction of the Mahler operator from $\mathbb{K}(x)$ to itself, which is not an automorphism of the field of rational functions. When the context is clear, we write D instead of (D, σ) .

Given a difference ring (D, σ) , a difference ring extension is a difference ring (D', σ') such that D' is an extension ring of D and σ' restricted to D is equal to σ . In practice, we will always write again σ for the extended σ' . The ring of constants of a difference ring D , denoted D^σ , is the subring of elements in D fixed by the endomorphism σ . If D is a difference field, the ring of constants is a field.

In what follows, every difference ring will be a difference ring extension of the field $\mathbb{K}(x)$ of rational functions. For the needed level of generality, let us introduce

$$\begin{aligned} (\mathbf{L}_\sigma) \quad & \ell_r \sigma^r y + \cdots + \ell_1 \sigma y + \ell_0 y = 0, \\ (\mathbf{R}_\sigma) \quad & \ell_r u \cdots \sigma^{r-1} u + \cdots + \ell_2 u \sigma u + \ell_1 u + \ell_0 = 0, \end{aligned}$$

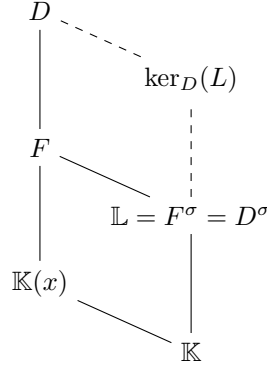


FIGURE 1. The inclusion relations for a difference field extension F of $\mathbb{K}(x)$ and a 1-universal ring extension D of it. The two rings F and D share the same field of constants \mathbb{L} . All inclusion but the dashed ones are difference ring inclusions; the dashed ones are only inclusions of \mathbb{L} -vector spaces. The space $\ker_D(L)$ is the space of the solutions of the operator L underlying the equation (L_σ) .

which we will refer to as, respectively, the linear difference equation (L_σ) and the Riccati difference equation (R_σ) .

2.2. First-order factors and their solutions. We now explain the link between the linear and the Riccati equations: the solutions to the Riccati equation are the coefficients u of the monic first-order right-hand factors $\sigma - u$ of the linear difference operator L .

Lemma 2.1. *Given an element u of a difference field extension F of $\mathbb{K}(x)$, the operator*

$$L = \ell_r(x)\sigma^r + \cdots + \ell_0(x) \in \mathbb{K}(x)\langle\sigma\rangle$$

associated with the linear difference equation (L_σ) admits $\sigma - u$ as a first-order right-hand factor in $F\langle\sigma\rangle$ if and only if u satisfies the Riccati difference equation (R_σ) .

Proof. The ring $F\langle\sigma\rangle$ is a skew Euclidean ring (Bronstein and Petkovšek 1993). By Euclidean division of L on the right by $\sigma - u$ in the ring $F\langle\sigma\rangle$, we obtain $L = \tilde{L}(\sigma - u) + R$ where R is exactly the left-hand side of (R_σ) . \square

We define hypergeometric elements in analogy with the classical difference case.

Definition 2.2. *Given a difference field extension F of $\mathbb{K}(x)$ and a difference ring extension D of F , an element y of D is F -hypergeometric if there exists $u \in F$ satisfying $\sigma y = uy$.*

Note that y may be zero in the definition, but that we will focus on nonzero hypergeometric y throughout.

The set of F -hypergeometric elements is generally not stable under addition. On the other hand, it is what we will call an F -cone: a nonempty set stable under multiplication by elements of F .

The following definition captures the notion of an extension of some difference field F that contains “enough” F -hypergeometric solutions of the linear equation (L_σ)

to suitably describe the “full” solution set of the Riccati equation (R_σ) in F . The definition is illustrated in Figure 1.

Definition 2.3. *Let F be a difference field extension of $\mathbb{K}(x)$, with constant field \mathbb{L} containing \mathbb{K} . A difference ring extension D of F is 1-universal if*

- (1) *the constant ring of D is the field \mathbb{L} ,*
- (2) *for each nonzero $u \in F$, the equation $\sigma y = uy$ has nonzero solutions in D .*

The extension D is said to have Property (U) if additionally

- (3) *for any linear difference equation (L_σ) of order r with coefficients in $\mathbb{K}(x)$, the vector space of solutions in D of this equation has dimension at most r over \mathbb{L} .*

In contrast with the classical theory, we do not consider any universal Picard–Vessiot algebra (van der Put and Singer 1997, Prop. 1.33; Roques 2018, Theorem 35): by point (2), only equations of order 1 are known to have solutions in D . On the other hand, classical hypotheses provide point (3), as shown by the following lemma.

Lemma 2.4. *Let $D \supset F \supset \mathbb{K}(x)$ be difference rings satisfying points (1) and (2) of Definition 2.3 for a suitable constant field \mathbb{L} containing \mathbb{K} . Assume that D is additionally simple, meaning that σ is an automorphism and that the only ideals of D stable under σ are (0) and D . Then, the extension D has Property (U).*

Proof. Under the classical assumption of simplicity (van der Put and Singer 1997, Def. 1.1; Singer 2016, Def. 4.1 and 4.4), the ring D^σ of constants is a field (van der Put and Singer 1997, Lemma 1.7). By our assumption $\ell_0 \ell_r \neq 0$ made in the introduction, the companion matrix A of L is invertible over $\mathbb{K}(x)$ and hence nonsingular over D . Suppose y_1, \dots, y_m are \mathbb{L} -linearly independent solutions of $Ly = 0$, and, for $1 \leq i \leq m$, introduce $Y_i = (y_i, \sigma y_i, \dots, \sigma^{r-1} y_i)^T$, which satisfies $\sigma Y_i = AY_i$. By point (1) and (Singer 2016, Lemma 4.8), this family $\{Y_i\}_{i=1}^m$ of solutions of $\sigma Y = AY$, which is linearly independent over \mathbb{L} , is linearly independent over D in the free D -module D^r . By the commutativity of D , the relation $m \leq r$ holds, thus proving Property (U). \square

The ring \mathfrak{D} that we will introduce for the Mahler case in §3 is a simple difference ring, as proved in an intermediate step of the proof of (Roques 2018, Théorème 35, top of page 342), where \mathfrak{D} is named B and is defined in (Roques 2018, Lemma 36). However, 1-universality and Property (U) are sufficient for our purposes and easily verified. We have preferred to follow this self-contained line of proof that relies on the tool of the Newton polygon, in the spirit of our previous article (CDDM 2018).

2.3. Parametrization of the solution set. Let F be a fixed difference field extension of $\mathbb{K}(x)$. Note that a 1-universal extension D of F is only a ring, so $\sigma y/y$ is not defined for all $y \in D$. However, restricting to nonzero F -hypergeometric y makes $1/y$ well defined, as stated in the following lemma.

Lemma 2.5. *Any nonzero F -hypergeometric element y of some 1-universal difference ring extension D over F has an F -hypergeometric inverse in D .*

Proof. For any nonzero F -hypergeometric $y \in D$, introduce $u \in F$ satisfying $\sigma y = uy$. From $y \neq 0$ follows $\sigma y \neq 0$, then $u \neq 0$. Then, also $1/u$ is in F , and $\sigma z = (1/u)z$ can be solved for a nonzero z , owing to the 1-universal nature of D over F . Then $\sigma(yz) = yz$ is a constant, which is in \mathbb{L} again because D is 1-universal over F . We find an inverse λz of y for an appropriate $\lambda \in \mathbb{L}$. \square

We now define similarity, a key concept in what follows.

Definition 2.6. *Given a difference field extension F of $\mathbb{K}(x)$ and a difference ring extension D of F , two elements y_1 and y_2 of D are F -similar if there exists a nonzero $q \in F$ satisfying $y_1 = qy_2$.*

Similarity is an equivalence relation, and y_1 and y_2 in the definition are either both nonzero or both zero. In what follows, we focus on nontrivial equivalence classes, that is, other than $\{0\}$, which, recalling our notation given in the introduction, we denote $C_{\neq 0} := C \setminus \{0\}$ for some set C containing 0.

Lemma 2.7. *For any F -similarity class $S_{\neq 0}$ of F -hypergeometric elements, the set S is an \mathbb{L} -vector space.*

Proof. Because \mathbb{L} is a subfield of F , S is an \mathbb{L} -cone. We show that it is also stable under addition. Consider two elements y_1 and y_2 in $S_{\neq 0}$ such that $y_1 + y_2 \neq 0$, and u and q in $F_{\neq 0}$ such that $\sigma y_1 = uy_1$ and $y_2 = qy_1$. Then, as $q \neq -1 \neq \sigma q$, $\sigma(y_1 + y_2) = (1 + \sigma q)uy_1 = (1 + \sigma q)(1 + q)^{-1}u(y_1 + y_2)$, from which follows that $y_1 + y_2$ is F -hypergeometric and F -similar to y_1 , thus in $S_{\neq 0}$. \square

We use the notation $\mathbb{P}(V)$ to denote the projectivization of a vector space V , leaving implicit the field over which V is defined. We also write $\mathbb{P}^n(K) = \mathbb{P}(K^{n+1})$ for the K -projective space of dimension n over a field K .

Theorem 2.8. *Let D be a 1-universal difference ring extension of F . For a fixed (\mathbb{L}_σ) , the following holds:*

- (1) *The \mathbb{L} -cone of F -hypergeometric solutions of (\mathbb{L}_σ) in D naturally partitions into the class $\{0\}$ and nontrivial F -similarity classes $\mathfrak{H}_{\neq 0}$. Moreover, the \mathbb{L} -cones \mathfrak{H} are \mathbb{L} -vector spaces in direct sum, and, under Property (U), their dimensions add up to at most r .*
- (2) *The set of solutions in F of (\mathbb{R}_σ) partitions into the images under the map $y \mapsto \sigma y/y$ of all nontrivial F -similarity classes of the set of F -hypergeometric solutions of (\mathbb{L}_σ) .*
- (3) *Each nontrivial F -similarity class $\mathfrak{H}_{\neq 0} \subset D$ of the set of F -hypergeometric solutions of (\mathbb{L}_σ) induces a one-to-one parametrization of its image under $y \mapsto \sigma y/y$ by the \mathbb{L} -projective space $\mathbb{P}(\mathfrak{H})$.*
- (4) *Assuming that D has Property (U), the sum of the dimensions of the parametrizing \mathbb{L} -projective spaces $\mathbb{P}(\mathfrak{H})$ does not exceed the order r of the equation (\mathbb{L}_σ) minus the number of nontrivial similarity classes $\mathfrak{H}_{\neq 0}$.*

Proof. (1) For each F -similarity class $\mathfrak{H}_{\neq 0}$ of F -hypergeometric solutions of L , Lemma 2.7 implies that \mathfrak{H} is an \mathbb{L} -vector space. We now prove that all the \mathfrak{H} are in direct sum. Given a relation $\lambda_0 y_0 + \cdots + \lambda_N y_N = 0$ with coefficients $\lambda_k \in \mathbb{L}_{\neq 0}$ between nonzero F -hypergeometric solutions y_k in distinct F -similarity classes, introduce $u_k \in F$ such that $\sigma y_k = u_k y_k$ for $0 \leq k \leq N$. Applying $\sigma - u_0$ yields

$$(2.1) \quad \lambda_1(u_1 - u_0)y_1 + \cdots + \lambda_N(u_N - u_0)y_N = 0.$$

Because the y_k were taken from distinct classes, none of the y_0/y_k , which exist by Lemma 2.5, can be a constant. So, none of the coefficients $\lambda_k(u_k - u_0)$ for $1 \leq k \leq N$ is zero, and (2.1) is a shorter nontrivial relation. Iterating the process, we obtain that some y_k is zero, a contradiction. As a consequence, the \mathbb{L} -vector spaces \mathfrak{H} are in direct sum. If D also satisfies Property (U) over F , this direct sum is included in an \mathbb{L} -space of dimension at most r , hence the dimension bound.

(2) Given a solution $u \in F$ of (R_σ) , Lemma 2.1 implies that $\sigma - u$ is a right-hand factor of L . Because D is 1-universal over F , there exists $y \neq 0$ such that $\sigma y = uy$, and y is invertible by Lemma 2.5. As a consequence, y is also a solution of (L_σ) and u equals $\sigma y/y$. So the map $y \mapsto \sigma y/y$ is a (well-defined) surjection from the set of nonzero F -hypergeometric solutions of (L_σ) to the set of solutions of (R_σ) in F . If u can be obtained by two solutions y_1 and y_2 , that is, if $\sigma y_1/y_1 = \sigma y_2/y_2$, then $\sigma(y_1/y_2) = y_1/y_2$ is a constant in D , thus in \mathbb{L} , so that, y_1 and y_2 are F -similar. As a consequence, the partitioning of the set of nonzero F -hypergeometric solutions of (L_σ) translates into a partitioning of the set of solutions of (R_σ) in F .

(3) By the proof of (2), a nontrivial F -similarity class $\mathfrak{H}_{\neq 0}$ of the set of F -hypergeometric solutions of (L_σ) maps under $y \mapsto \sigma y/y$ to a subset U of the set of solutions of (R_σ) in F . Given y_1 and y_2 in the same class $\mathfrak{H}_{\neq 0}$, thus satisfying $y_1 = qy_2$ for some $q \in F$ (which needs to be nonzero), if both share the same image in U , that is, $\sigma y_1/y_1 = \sigma y_2/y_2$ is the same element of U , then the proof of (2) has shown that q is a constant from \mathbb{L} . This implies that y_1 and y_2 are in the same \mathbb{L} -projective class, thus proving the result.

(4) Property (U) implies that each similarity class \mathfrak{H} is finite-dimensional. If it has \mathbb{L} -dimension s , its associated \mathbb{L} -projective space $\mathbb{P}(\mathfrak{H})$ has dimension $s - 1$. The statement is next a direct consequence of points (1) and (3). \square

Remark 2.9. Under Property (U), point (3) of Theorem 2.8 leads to a more explicit parametrization, owing to the finite \mathbb{L} -dimension of the space \mathfrak{H} . Fix y_0 in $\mathfrak{H}_{\neq 0}$ and let u_0 satisfy $\sigma y_0 = u_0 y_0$. Let $(q_1 y_0, \dots, q_s y_0)$ be some basis of \mathfrak{H} . With coordinates, the bijection of the theorem translates into the map

$$(c_1 : \dots : c_s) \mapsto \frac{\sigma(c_1 q_1 + \dots + c_s q_s)}{c_1 q_1 + \dots + c_s q_s} u_0,$$

now a rational bijection from $\mathbb{P}^{s-1}(\mathbb{L})$ to its image.

Remark 2.10. The partition obtained by point (2) of the theorem can be interpreted as the partition induced on the solution set of (R_σ) by the equivalence relation on $F_{\neq 0}$ defined by: u and u' are F -equivalent if there exists $f \in F_{\neq 0}$ such that $u/u' = \sigma f/f$.

3. GENERALIZED SERIES SOLUTIONS OF THE LINEAR EQUATION

In this section, we define a space, denoted \mathfrak{D} , of explicit formal expressions that is a 1-universal extension with Property (U) of both $\overline{\mathbb{K}}(x^{1/*})$ and $\overline{\mathbb{K}}((x^{1/*}))$. This will be a crucial ingredient to describe the Puiseux series solutions of a Riccati Mahler equation (Theorem 3.14). To this end, we describe $\mathbb{L}((x^{1/*}))$ -hypergeometric elements (Proposition 3.12) so as to apply Theorem 2.8 for a given field \mathbb{L} . The properties of \mathfrak{D} are again used in §4 to describe the ramified rational solutions of the Riccati equation (Theorem 4.5).

3.1. The difference ring \mathfrak{D} . Slightly adapting (Roques 2018, §5.2), we introduce the $\overline{\mathbb{K}}((x^{1/*}))$ -algebra \mathfrak{D} generated by elements e_λ , $\lambda \in \overline{\mathbb{K}}_{\neq 0}$, satisfying the algebraic relations $e_\lambda e_{\lambda'} = e_{\lambda\lambda'}$ and $e_1 = 1$, so as to have

$$\mathfrak{D} = \bigoplus_{\lambda \in \overline{\mathbb{K}}_{\neq 0}} e_\lambda \overline{\mathbb{K}}((x^{1/*})).$$

This ring is not a domain, as shown by the product $(e_\lambda + e_{-\lambda})(e_\lambda - e_{-\lambda}) = 0$. Next, we equip the algebra \mathfrak{D} with a structure of difference ring by enforcing $Me_\lambda = \lambda e_\lambda$.

To support the intuition, we henceforth denote e_λ by $(\ln x)^{\log_b \lambda}$ for $\lambda \in \overline{\mathbb{K}}_{\neq 0}$. This mere notation bears no analytic meaning, but it is reminiscent of the relation $M \ln(x) = \ln(x^b) = b \ln(x)$ of its analytic counterpart, as well as of the linear independence over the field of meromorphic functions of the family of functions $((\ln x)^{\log_b \lambda})_{\lambda \in \mathbb{C}_{\neq 0}}$. As a consequence the difference ring is expressed as

$$(3.1) \quad \mathfrak{D} = \bigoplus_{\lambda \in \overline{\mathbb{K}}_{\neq 0}} (\ln x)^{\log_b \lambda} \overline{\mathbb{K}}((x^{1/*})),$$

where $(\ln x)^{\log_b \lambda}$ is an eigenvector of M with respect to the eigenvalue λ .

For an operator of $\overline{\mathbb{K}}(x)\langle M \rangle$,

$$(3.2) \quad L = L(x, M) = \sum_{k=0}^r \ell_k(x) M^k,$$

and an element of \mathfrak{D} ,

$$(3.3) \quad y = \sum_{\lambda \in \Lambda} (\ln x)^{\log_b \lambda} p_\lambda(x),$$

where the set Λ is a finite subset of $\overline{\mathbb{K}}$ and the $p_\lambda(x)$ are Puiseux series, the action of L on y has a special structure given by

$$(3.4) \quad L(x, M) y = \sum_{\lambda \in \Lambda} (\ln x)^{\log_b \lambda} L(x, \lambda M) p_\lambda(x).$$

This formula decomposes (L) into equations on the $p_\lambda(x)$ as follows.

Lemma 3.1. *For an operator $L \in \overline{\mathbb{K}}(x)\langle M \rangle$ as in (3.2) and an element y of \mathfrak{D} as in (3.3), the equality $L(x, M) y = 0$ is equivalent to the finite set of equalities*

$$(3.5) \quad L(x, \lambda M) p_\lambda(x) = 0, \quad \lambda \in \Lambda.$$

Proof. This follows directly from (3.4), by the uniqueness of coefficients in \mathfrak{D} . \square

Given a subfield \mathbb{L} of $\overline{\mathbb{K}}$, we also define a subring $\mathfrak{D}_{\mathbb{L}}$ of \mathfrak{D} by

$$(3.6) \quad \mathfrak{D}_{\mathbb{L}} = \bigoplus_{\lambda \in \mathbb{L}_{\neq 0}} (\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/*})).$$

It is obviously a difference ring and \mathfrak{D} is the special case $\mathfrak{D}_{\overline{\mathbb{K}}}$. The following lemma can be deduced from Roques (2018, Theorem 35) in the case $\mathbb{L} = \overline{\mathbb{Q}}$.

Lemma 3.2. *The field of constants of the ring $\mathfrak{D}_{\mathbb{L}}$ is \mathbb{L} .*

Proof. By Lemma 3.1, the equation $My = y$ is equivalent to $\lambda M p_\lambda = p_\lambda$ for $\lambda \in \Lambda$. For each λ , if p_λ is nonzero, its valuation must be 0 and its leading coefficient c must satisfy $\lambda c = c$, implying $\lambda = 1$. This means that y is in \mathbb{L} . We conclude that the constant field of $\mathfrak{D}_{\mathbb{L}}$ is \mathbb{L} itself. \square

3.2. Bounds. The goal of this section is to bound the valuation and ramification order of the Puiseux series solutions of (\mathbf{L}) in \mathfrak{D} . To the polynomial

$$L = \sum_{k=0}^r \ell_k(x) M^k = \sum_{k,j} \ell_{k,j} x^j M^k$$

we associate the set of points $(b^k, j) \in \mathbb{R}^2$ such that $\ell_{k,j} \neq 0$. The lower convex hull of this set is called the (lower) Newton polygon (of L). The following lemma is borrowed from (CDDM 2018, Lemma 2.2).

Lemma 3.3. *The valuation of any Puiseux series solution of (\mathbf{L}) is the opposite of the slope of an edge of the lower Newton polygon of L . Moreover the coefficients $\ell_{k,j}$ corresponding to all points (b^k, j) lying on the edge must add up to zero.*

For a solution of (\mathbf{L}) , that is, satisfying $L(x, M)y = 0$, to be of the form $y = (\ln x)^{\log_b \lambda} p(x)$ with $p(x)$ a Puiseux series, the series $p(x)$ must by the formula (3.4) be a solution of $L(x, \lambda M)p = 0$. The change from $L(x, M)$ to $L(x, \lambda M)$ does not modify the Newton polygon, but the coefficients at abscissa b^k are multiplied by λ^k . So, by Lemma 3.3, the only λ for which solutions of the form $(\ln x)^{\log_b \lambda} p(x)$ may exist are therefore the (necessary nonzero) roots of the polynomial $\chi(X) := \sum_{k_1}^{k_2} \ell_{k,j} X^{k-k_1}$ where b^{k_1} and b^{k_2} are the abscissae of the endpoints of the edge, with $k_1 \leq k_2$. We call χ the *characteristic polynomial* of the edge.

Next, we index the edges of the polygon so that their slopes form an increasing sequence. We denote by χ_j the characteristic polynomial of the j th edge and we define Λ as the union of the roots of these polynomials:

$$(3.7) \quad \Lambda = \{\lambda \in \overline{\mathbb{K}} : \exists j, \chi_j(\lambda) = 0\}.$$

Now fix λ in Λ . The j th edge will be called λ -*admissible* (for L) if $\chi_j(\lambda) = 0$. Note that 1-admissible edges are called *admissible* in (CDDM 2018). Lemma 3.3 extends slightly to accommodate a logarithmic part, leading to the following result. The omitted proof parallels that of (CDDM 2018, Lemma 2.2).

Lemma 3.4. *The valuation of any solution of (\mathbf{L}) of the form $(\ln x)^{\log_b \lambda} p(x)$ for some nonzero Puiseux series $p(x)$ is the opposite of the slope of a λ -admissible edge of the lower Newton polygon of L .*

As already noted for $\lambda = 1$ in (CDDM 2018), λ -admissibility is only a necessary condition for the existence of a series solution with logarithmic part $(\ln x)^{\log_b \lambda}$.

In order to discuss ramifications, we import (CDDM 2018, Prop. 2.19), which deals with 1-admissible edges. In doing so, we slightly reformulate the result so as to describe solutions over an algebraic extension \mathbb{L} of \mathbb{K} .

Lemma 3.5. *Any Puiseux series solution of (\mathbf{L}) is an element of $\overline{\mathbb{K}}((x^{1/q_1}))$, where q_1 is the lcm of the denominators of the slopes of those 1-admissible edges such that said denominators are coprime with b . Furthermore, a $\overline{\mathbb{K}}$ -vector basis of the solutions can be written in $\mathbb{K}((x^{1/q_1}))$. Given an intermediate field \mathbb{L} , such a $\overline{\mathbb{K}}$ -vector basis is also an \mathbb{L} -vector basis for the \mathbb{L} -space of solutions in $\mathbb{L}((x^{1/*}))$.*

Proof. The first assertion of the statement is exactly the proposition we referred to, using $\overline{\mathbb{K}}$ as a base field. The second and third assertions are implicit in (CDDM 2018), which finds solutions by solving linear systems over \mathbb{K} . \square

More generally, we denote by q_λ the lcm of the denominators coprime with b of the slopes of the λ -admissible edges of the lower Newton polygon of L , that is

$$(3.8) \quad q_\lambda = \text{lcm} \{q \in \mathbb{N} : \exists j, p, \gcd(b, q) = \gcd(p, q) = 1, \chi_j(\lambda) = 0, \\ \text{and the } j\text{th edge has slope } p/q\}.$$

We can then prove the following result.

Proposition 3.6. *Given an intermediate field \mathbb{L} , the solution space of (L) in $\mathfrak{D}_{\mathbb{L}}$ is included in*

$$\bigoplus_{\lambda \in \Lambda \cap \mathbb{L}} (\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/q_\lambda})).$$

Furthermore, an \mathbb{L} -basis of this space can be written in

$$\prod_{\lambda \in \Lambda \cap \mathbb{L}} (\ln x)^{\log_b \lambda} \mathbb{K}[\lambda]((x^{1/q_\lambda})).$$

Proof. For a nonzero solution y of L in $\mathfrak{D}_{\mathbb{L}}$, Lemma 3.1 splits the relation $Ly = 0$ into the equivalent finite set of equations (3.5). This proves the first part, where the direct sum of the solution space is induced by the direct sum in the expression (3.6) of $\mathfrak{D}_{\mathbb{L}}$. Next, according to Lemma 3.3, the equation (3.5) for a given $\lambda \in \mathbb{L}$ has a nonzero solution only if there exists a λ -admissible edge, that is, only if λ is an element of Λ . Moreover the change from $L(x, M)$ to $L(x, \lambda M)$ modifies the linear system that determines the solutions by moving its coefficients to the algebraic extension $\mathbb{K}[\lambda]$, so that, by Lemma 3.5, a basis of their Puiseux-series solutions p_λ can be found with coefficients in $\mathbb{K}[\lambda]$ and ramification order q_λ . The result is proved. \square

Corollary 3.7. *Given an intermediate field \mathbb{L} , the solution set in $\mathfrak{D}_{\mathbb{L}}$ of the linear equation (L) is an \mathbb{L} -vector space whose dimension over the field \mathbb{L} is at most the order r of the equation.*

Proof. Proposition 3.6 provides us with an \mathbb{L} -basis of the solutions of (L) in $\mathfrak{D}_{\mathbb{L}}$, which can be chosen such that for each λ the basis elements of the form $(\ln x)^{\log_b \lambda} p(x)$ have distinct valuations. By Lemma 3.4, these valuations are opposites of slopes of distinct edges of the Newton polygon whose characteristic polynomials χ_j all have λ as a root. Therefore, the total number of elements of the basis is bounded by $\sum_j \deg \chi_j \leq r$. \square

The order r is not always reached in Corollary 3.7, as shown by the following two examples.

Example 3.8. When $b = 2$, the operator $L = M^2 - 2M + 1$ admits the series solution 1, and no nonconstant solution. This reflects the fact that the nonhomogeneous equation $My - y = 1$ has no solution in \mathfrak{D} , while any solution of this equation needs to be killed by L .

Example 3.9. Similarly, when $b = 2$, $xM^2 - (1 + x)M + 1$ admits the series solution 1, and otherwise would require a Hahn series $x^{-1/2} + x^{-1/4} + x^{-1/8} + \dots$ to get a second dimension of solutions, the latter solving $My - y = 1/x$.

3.3. Puiseux-hypergeometric solutions. In the present section, we focus on the field $F = \mathbb{L}((x^{1/*}))$ to obtain the Puiseux series with coefficients in \mathbb{L} that are solutions of the Riccati equation. We describe the $\mathbb{L}((x^{1/*}))$ -hypergeometric elements in $\mathfrak{D}_{\mathbb{L}}$ and deduce that $\mathfrak{D}_{\mathbb{L}}$ is a 1-universal extension of $\mathbb{L}((x^{1/*}))$.

Lemma 3.10. *Fix $q \in \mathbb{N}_{\neq 0}$ and $u \in \mathbb{L}((x^{1/q}))$ of valuation α and with leading coefficient 1. Then, the equation $Mv = uv$ admits a unique series solution v in $x^{\alpha/(b-1)}\mathbb{L}[[x^{1/q}]]$ of valuation $\alpha/(b-1)$ and with leading coefficient 1. The solution space of this equation in $\mathbb{L}((x^{1/*}))$ is exactly the one-dimensional space $\mathbb{L}v$.*

Proof. The Puiseux series u can be written in the form $u = x^{\alpha}\hat{u}$ for some $\hat{u} = \sum_{k \in \mathbb{N}} \hat{u}_k x^{k/q} \in \mathbb{L}[[x^{1/q}]]$ such that $\hat{u}_0 = 1$. The announced valuation of v is a necessary consequence of the equation $Mv = uv$. Given v of the form $x^{\alpha/(b-1)}\hat{v}$ $\hat{v} = \sum_{k \in \mathbb{N}} \hat{v}_k x^{k/q}$ and $\hat{v}_0 \neq 0$, the equation is equivalent to the relation

$$\sum_{k \in \mathbb{N}} \hat{v}_k x^{bk/q} = \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{N}} \hat{u}_{n-i} \hat{v}_i x^{n/q},$$

which in turn is equivalent to the recurrence formula

$$\hat{v}_n = \hat{v}_{n/b} - \sum_{i=0}^{n-1} \hat{u}_{n-i} \hat{v}_i$$

for $n \in \mathbb{N}_{\neq 0}$, where $\hat{v}_{n/b}$ is understood to be zero if n/b is not in \mathbb{N} . The sequence $(\hat{v}_n)_{n \in \mathbb{N}}$, and therefore the solution v , is fully determined by the choice of $\hat{v}_0 = 1$.

Next, for any Puiseux series solution w of the equation $Mw = uw$, the quotient w/v is in the constant field of $\mathbb{L}((x^{1/*}))$, that is, it lies in \mathbb{L} . \square

Corollary 3.11. *$\mathfrak{D}_{\mathbb{L}}$ is a 1-universal extension with Property (U) of any intermediate difference field F between $\mathbb{L}(x)$ and $\mathbb{L}((x^{1/*}))$, so that Theorem 2.8 applies to $\mathfrak{D}_{\mathbb{L}}$, viewed as an extension of F .*

Proof. This follows directly from Lemma 3.2, Corollary 3.7, and Lemma 3.10. \square

We will use $F = \mathbb{L}((x^{1/*}))$ in Theorem 3.14 below, and $\mathbb{L}(x^{1/*})$ in Theorem 4.5.

Proposition 3.12. *The $\mathbb{L}((x^{1/*}))$ -hypergeometric elements in $\mathfrak{D}_{\mathbb{L}}$ are its elements $y = (\ln x)^{\log_b \lambda} p(x)$ where λ is in $\mathbb{L}_{\neq 0}$ and $p(x)$ is a Puiseux series in $\mathbb{L}((x^{1/*}))$. Two such nonzero elements are similar if and only if they share the same λ .*

Proof. Let y be a nonzero $\mathbb{L}((x^{1/*}))$ -hypergeometric element in $\mathfrak{D}_{\mathbb{L}}$, together with the corresponding Puiseux series u satisfying the equation $My = uy$. The series u lies in a subfield $\mathbb{L}((x^{1/q}))$ with q a positive integer. Let its leading term be denoted $\lambda_0 x^{m/q}$, for suitable $\lambda_0 \in \mathbb{L}$ and $m \in \mathbb{Z}$. By Lemma 3.1, writing y as in (3.3) for a finite subset Λ of $\mathbb{L}_{\neq 0}$ and some $p_{\lambda}(x)$ in $\mathbb{L}((x^{1/*}))$ leads to the equations $\lambda M p_{\lambda} = u p_{\lambda}$ for $\lambda \in \Lambda$. Any such equation can have a nonzero solution only if λ is equal to λ_0 , forcing y to be of the form $y = (\ln x)^{\log_b \lambda_0} p_{\lambda_0}(x)$ where $p_{\lambda_0}(x) \in \mathbb{L}((x^{1/*}))$ is nonzero. Conversely, any y of that form is $\mathbb{L}((x^{1/*}))$ -hypergeometric in $\mathfrak{D}_{\mathbb{L}}$.

Now suppose that $y_1 = (\ln x)^{\log_b \lambda_1} p_1(x)$ and $y_2 = (\ln x)^{\log_b \lambda_2} p_2(x)$ are $\mathbb{L}((x^{1/*}))$ -similar and nonzero. The similarity implies the linear dependence of $(\ln x)^{\log_b \lambda_1}$ and $(\ln x)^{\log_b \lambda_2}$ over $\mathbb{L}((x^{1/*}))$, and therefore the equality $\lambda_1 = \lambda_2$. Conversely two nonzero hypergeometric solutions of (L) in $\mathfrak{D}_{\mathbb{L}}$ with the same $\lambda \in \mathbb{L} \cap \Lambda'_*$ are similar by definition. \square

Remark 3.13. The same approach, supplemented with Lemma 3.10, shows more precisely that the elements $\mathbb{L}((x^{1/q}))$ -hypergeometric in $\mathfrak{D}_{\mathbb{L}}$ write

$$(3.9) \quad y = (\ln x)^{\log_b \lambda} x^{m/(q(b-1))} s(x^{1/q})$$

with λ in $\mathbb{L}_{\neq 0}$, q a positive integer, m an integer, and $s(x)$ a series of valuation 0 in $\mathbb{L}[[x]]$. Observe (e.g., making $m = 1$) that My/y may have a lower ramification order than y . Consequently, obtaining *all* Laurent power series solutions to (R) in the form $u = My/y$ generally requires to consider *ramified* hypergeometric solutions y of (L). A simple example is provided by $L = M - u$ for $b = 5$ and $u = x(1 + x^3)$, with solution $y = x^{1/4} \prod_{k \geq 0} (1 + x^{3 \cdot 5^k})^{-1}$.

Let Λ' denote the finite set of $\lambda \in \Lambda$ such that there exists a nonzero solution of (L) in $(\ln x)^{\log_b \lambda} \overline{\mathbb{K}}((x^{1/*}))$. For $\lambda \in \Lambda'$, define \mathfrak{H}_{λ} to be the set of hypergeometric solutions in $(\ln x)^{\log_b \lambda} \overline{\mathbb{K}}((x^{1/*}))$, which are necessarily in $(\ln x)^{\log_b \lambda} \overline{\mathbb{K}}((x^{1/q\lambda}))$ by Proposition 3.6. With this notation, Theorem 2.8 specializes as follows.

Theorem 3.14. *The solution set $\mathfrak{R}_{\mathbb{L}((x^{1/*}))}$ of the Riccati equation (R) in $\mathbb{L}((x^{1/*}))$ is a disjoint union indexed by the finite set $\mathbb{L} \cap \Lambda'$,*

$$(3.10) \quad \mathfrak{R}_{\mathbb{L}((x^{1/*}))} = \coprod_{\lambda \in \mathbb{L} \cap \Lambda'} \mathfrak{R}_{\mathbb{L}, \lambda}.$$

Each set $\mathfrak{R}_{\mathbb{L}, \lambda}$ is a set of series from $\mathbb{L}((x^{1/q\lambda}))$ with leading coefficient λ and is one-to-one rationally parametrized by the \mathbb{L} -projective space $\mathbb{P}(\mathfrak{H}_{\lambda} \cap \mathfrak{D}_{\mathbb{L}})$. The corresponding parametrization is obtained by restricting the map $y \mapsto My/y$ from $(\mathfrak{H}_{\lambda})_{\neq 0} \cap \mathfrak{D}_{\mathbb{L}}$ to its image in $\mathbb{L}((x^{1/*}))$ and projectivizing its source. Moreover the dimensions of the \mathbb{L} -projective spaces add up to a number that is at most the order r of the linear equation (L) minus the cardinality of $\mathbb{L} \cap \Lambda'$.

Proof. By Corollary 3.11, the ring $\mathfrak{D}_{\mathbb{L}}$ has Property (U) as an extension of $\mathbb{L}((x^{1/*}))$, so Theorem 2.8 applies to $D = \mathfrak{D}_{\mathbb{L}}$ and $F = \mathbb{L}((x^{1/*}))$. Points (1) to (3) of that theorem justify the partitioning according to (3.10) as well as the form of the parametrization. Moreover, from the inclusion $\mathfrak{H}_{\lambda} \subset (\ln x)^{\log_b \lambda} \overline{\mathbb{K}}((x^{1/q\lambda}))$ follows the inclusion $\mathfrak{R}_{\mathbb{L}, \lambda} \subset \mathbb{L}((x^{1/q\lambda}))$. Point (4) proves the bound on dimensions. \square

4. RAMIFIED RATIONAL SOLUTIONS TO THE RICCATI EQUATION

We now study the F -hypergeometric solutions of (L) for $F = \mathbb{L}(x^{1/*})$, or, equivalently, the ramified rational solutions of (R) with coefficients in \mathbb{L} .

4.1. Hypergeometric elements. We begin by characterizing the $\mathbb{L}(x^{1/*})$ -hypergeometric elements of $\mathfrak{D}_{\mathbb{L}}$.

Proposition 4.1. *The nonzero $\mathbb{L}(x^{1/*})$ -hypergeometric elements in $\mathfrak{D}_{\mathbb{L}}$ are the elements*

$$(4.1) \quad y = (\ln x)^{\log_b \lambda} x^{m/q} f(x^{1/q})$$

with λ in $\mathbb{L}_{\neq 0}$, m an integer, q a positive integer, and $f(x)$ an infinite product

$$(4.2) \quad f(x) = \prod_{k \geq 0} 1/g(x^{b^k})$$

given by $g(x)$ in $\mathbb{L}(x)$ satisfying $g(0) = 1$, which converges in the ring of formal power series $\mathbb{L}[[x]]$.

Proof. Any nonzero $\mathbb{L}(x^{1/*})$ -hypergeometric element y in $\mathfrak{D}_{\mathbb{L}}$ is $\mathbb{L}((x^{1/*}))$ -hypergeometric. By Proposition 3.12, it can be written $y = (\ln x)^{\log_b \lambda} x^{m/q} s(x^{1/q})$ with λ in $\mathbb{L}_{\neq 0}$, q a positive integer, m an integer, and $s(x)$ in $\mathbb{L}[[x]]$ of valuation 0. One then has $My/y = \lambda x^{(b-1)m/q} s(x^{b/q})/s(x^{1/q}) \in \mathbb{L}((x^{1/q})) \cap \mathbb{L}(x^{1/*})$, so $g(x) := s(x^b)/s(x)$ is a rational function. The infinite product (4.2) for this value of g converges to s , hence y is of the form (4.1) with $f = s$. \square

Remark 4.2. If the hypergeometric element y is such that My/y is in $\mathbb{L}((x^{1/q'}))$, then $q = q'(b-1)$ can be used in formula (4.1), as a consequence of Remark 3.13.

Remark 4.3. We could have already written the series s in Proposition 3.12 as an infinite product. But doing so proves useful only in the present situation, via the interpretation of g and the following characterization of the similarity classes in Proposition 4.4.

4.2. Similarity classes. The set of $\mathbb{L}(x^{1/*})$ -hypergeometric solutions of (L) decomposes into $\mathbb{L}(x^{1/*})$ -similarity classes that refine the $\mathbb{L}((x^{1/*}))$ -similarity classes described in §3. The following proposition gives a normal form for these similarity classes. Recall from Proposition 3.6 that the integer q_λ defined by (3.8) is a bound on the ramification orders of solutions of (L) in $(\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/*}))$.

Proposition 4.4. *Consider a nontrivial $\mathbb{L}(x^{1/*})$ -similarity class $\mathfrak{H}_{\neq 0}$ of $\mathbb{L}(x^{1/*})$ -hypergeometric solutions of (L). Then, there exist:*

- λ in $\mathbb{L} \cap \Lambda'$,
- an integer $s > 0$ and Laurent polynomials p_1, \dots, p_s of $\mathbb{L}[x, x^{-1}]$,
- a rational function g in $\mathbb{L}(x)$ satisfying $g(0) = 1$,

such that the generalized series

$$(4.3) \quad (\ln x)^{\log_b \lambda} \times p_j(x^{1/q_\lambda}) \times \prod_{k \geq 0} 1/g(x^{b^k/q_\lambda}), \quad 1 \leq j \leq s,$$

form a basis of the \mathbb{L} -vector space \mathfrak{H} .

One may additionally require that:

- (1) the family (p_1, \dots, p_s) is in reduced echelon form w.r.t. ascending degree (meaning in particular that the coefficient of minimal degree of p_i is 1),
- (2) the elements p_1, \dots, p_s are coprime in $\mathbb{L}[x, x^{-1}]$ (in other words, the $p_i/x^{\text{val } p_i}$ are coprime in $\mathbb{L}[x]$).

The tuple $(\lambda, g, p_1, \dots, p_s)$ is then uniquely determined.

Moreover, two bases (4.3) using pairs (λ, g) and (λ, g') with the same λ describe the same class if and only if the infinite product $\pi := \prod_{k \geq 0} M^k(g'/g)$ is a rational function.

Proof. The $\mathbb{L}(x^{1/*})$ -similarity class $\mathfrak{H}_{\neq 0}$ of $\mathbb{L}(x^{1/*})$ -hypergeometric solutions of (L) is contained in an $\mathbb{L}((x^{1/*}))$ -similarity class of $\mathbb{L}((x^{1/*}))$ -hypergeometric elements of $\mathfrak{D}_{\mathbb{L}}$, which, by Proposition 3.12, is of the form $(\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/*}))_{\neq 0}$ for some $\lambda \in \mathbb{L}_{\neq 0}$. Since, additionally, \mathfrak{H} is contained in the solution space of (L), Proposition 3.6 implies $\mathfrak{H} \subseteq (\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/q_\lambda}))$. Additionally, \mathfrak{H} has finite dimension over \mathbb{L} by Theorem 2.8.

For any element $y \in \mathfrak{H}_{\neq 0}$, use Proposition 4.1 to write

$$(4.4) \quad y = (\ln x)^{\log_b \lambda} x^{m/q} f(x^{1/q}), \quad f(x) = \prod_{k \geq 0} 1/g(x^{b^k}),$$

with $g(x) \in \mathbb{L}(x)$ and $g(0) = 1$. We contend that we can make $q = q_\lambda$ in (4.4) without loss of generality. First, the valuation m/q of $y/(\ln x)^{\log_b \lambda}$ is in $q_\lambda^{-1}\mathbb{Z}$, and can thus be written $m/q = \tilde{m}/q_\lambda$ for $\tilde{m} \in \mathbb{Z}$. Second, $f(x^{1/q}) = y/((\ln x)^{\log_b \lambda} x^{\tilde{m}/q_\lambda})$ is in $\mathbb{L}((x^{1/q_\lambda}))$, so that $h(x) := f(x^{q_\lambda/q})$ is in $\mathbb{L}((x))$. From $Mf/f = g \in \mathbb{L}(x)$ follows $g(x^{q_\lambda/q}) = Mh/h \in \mathbb{L}(x^{q_\lambda/q}) \cap \mathbb{L}((x))$. Write $q_\lambda/q = n/d$ in lowest terms and observe that g is a series in x^d . Define $\tilde{g}(x) := g(x^{n/d})$, a rational series satisfying $\tilde{g}(0) = 1$. Define $\tilde{f}(x) := \prod_{k \geq 0} 1/\tilde{g}(x^{b^k})$ and observe $f(x^{1/q}) = \tilde{f}(x^{1/q_\lambda})$. We have obtained a new expression of y of the form (4.4) with (q, m, g) replaced with $(q_\lambda, \tilde{m}, \tilde{g})$, and a ramification order constrained to be q_λ .

Pick an element $y_0 \in \mathfrak{H}_{\neq 0}$ and obtain its decomposition of the form (4.4), thus fixing (q, m, g) to some (q_λ, m_0, g_0) . Since $\mathfrak{H}_{\neq 0}$ is an $\mathbb{L}(x^{1/*})$ -similarity class, the set $V := \{y/y_0 : y \in \mathfrak{H}\}$ is an \mathbb{L} -subspace of $\mathbb{L}(x^{1/*})$, and thus, by the previous points, a finite-dimensional \mathbb{L} -subspace of $\mathbb{L}(x^{1/q_\lambda})$. Let $a(x^{1/q_\lambda})$ be the least common denominator of the elements of V , with valuation coefficient normalized to one. Write $a(x^{1/q_\lambda}) = x^{v/q_\lambda} \tilde{a}(x^{1/q_\lambda})$ for \tilde{a} of valuation zero. Then $a(x^{1/q_\lambda})V$ is a finite-dimensional subspace of $\mathbb{L}[x^{1/q_\lambda}]$; as such, it admits a basis $(\tilde{p}_1(x^{1/q_\lambda}), \dots, \tilde{p}_s(x^{1/q_\lambda}))$ in reduced echelon form w.r.t. ascending degree. The polynomials $\tilde{p}_1, \dots, \tilde{p}_s$ are coprime by the minimality of a . Setting $g = g_0 a / M a$ and $p_j = x^{m_0 - v} \tilde{p}_j$, the expressions (4.3) form a basis of \mathfrak{H} that satisfies the conditions (1) and (2).

We already have noticed that λ is uniquely determined by \mathfrak{H} . Suppose that \mathfrak{H} admits a second basis of the form (4.3), with parameters g', p'_1, \dots, p'_s also satisfying (1) and (2). Let $f' = \prod_{k \geq 0} 1/g'(x^{b^k})$. Then (p_1, \dots, p_s) and $(p'_1 f'/f, \dots, p'_s f'/f)$ are two bases of the same \mathbb{L} -vector space. In particular, the $p'_j f'/f$ are Laurent polynomials. A Bézout relation $\sum_j u_j p'_j = 1$ for Laurent polynomials u_j then implies that $f'/f = \sum_j u_j (p'_j f'/f)$ is a Laurent polynomial. By symmetry, f/f' is also Laurent, and in fact equal to 1 because $g(0) = g'(0) = 1$. Then the uniqueness of reduced echelon forms implies $(p_1, \dots, p_s) = (p'_1, \dots, p'_s)$. We have thus proved the uniqueness of tuple $(\lambda, g, p_1, \dots, p_s)$ under conditions (1) and (2).

Finally, suppose two classes are described by bases of the form (4.3) for pairs (λ, g) and (λ, g') with the same λ and suitable families of Laurent polynomials. Fix the quotient of a function from one class by a function from the other: this is the product of the infinite product $\pi(x^{1/q_\lambda})$ by a ramified rational function. Thus, the two classes coincide if and only if $\pi(x)$ is simultaneously a ramified rational function and a formal power series, in other words, if and only if it is a rational series, showing the last point of the proposition. \square

Summing up the previous discussion, the full solution set of the Riccati equation can be described as follows.

Theorem 4.5. *The solution set $\mathfrak{R}_{\mathbb{L}(x^{1/*})}$ of the Riccati equation (R) in $\mathbb{L}(x^{1/*})$ is a disjoint union*

$$(4.5) \quad \mathfrak{R}_{\mathbb{L}(x^{1/*})} = \coprod_{(\lambda, g)} \mathfrak{R}_{\mathbb{L}, \lambda, g}$$

indexed by a finite set of pairs (λ, g) , where λ is in $\mathbb{L} \cap \Lambda'$ and g is a rational function satisfying $g(0) = 1$, where each $\mathfrak{R}_{\mathbb{L}, \lambda, g}$ can be parametrized bijectively using

the notation of Proposition 4.4 by

$$u(x) = \lambda \frac{\sum_{j=1}^s c_j p_j(x^{b/q\lambda})}{\sum_{j=1}^s c_j p_j(x^{1/q\lambda})} g(x^{1/q\lambda}) \quad \text{for } (c_1 : \dots : c_s) \in \mathbb{P}^{s-1}(\mathbb{L}).$$

The dimensions $s - 1$ of the projective spaces add up to a number that is at most the order r of the linear equation (L) minus the cardinality of the set of pairs (λ, g) indexing the disjoint union (4.5).

Proof. We apply Theorem 2.8 in the light of Proposition 4.4. Given an $\mathbb{L}(x^{1/*})$ -similarity class $\mathfrak{H}_{\neq 0}$ of $\mathbb{L}(x^{1/*})$ -hypergeometric solutions of (L), Proposition 4.4 provides us with a pair (λ, g) that identifies it uniquely. Call $\mathfrak{R}_{\mathbb{L}, \lambda, g}$ its image under the map $y \mapsto My/y$. Then, Theorem 2.8 and Remark 2.9 provide the result. \square

5. DEGREE BOUNDS FOR RATIONAL SOLUTIONS

In the present section, we derive degree bounds for rational solutions of the Riccati equation in terms of its order r and coefficient degree d . Such bounds will be useful in §7.1 to select relevant Hermite–Padé approximants.

The following easy consequence of Lemma 3.1(c) in (CDDM 2018) will be used throughout; we prove it for self-containedness.

Lemma 5.1. *Given P and Q in $\overline{\mathbb{Q}}[x]$ and any $j \in \mathbb{N}$, P and Q are coprime if and only if so are $M^j P$ and $M^j Q$.*

Proof. If P and Q are coprime, there exist polynomials A and B such that $AP + BQ = 1$. Therefore, $M^j A M^j P + M^j B M^j Q = 1$, so that $M^j P$ and $M^j Q$ are coprime. Conversely, if $M^j P$ and $M^j Q$ are coprime, then there exist polynomials A and B such that $A M^j P + B M^j Q = 1$. If some $G \in \overline{\mathbb{Q}}[x]$ divides both P and Q , then $M^j G$ divides 1, so G must be a constant. Therefore, P and Q are coprime. \square

Proposition 5.2. *Assume that $P/Q \in \overline{\mathbb{K}}(x)$, with $\gcd(P, Q) = 1$, is a solution of (R). Then, the following degree bounds hold:*

$$(5.1) \quad \deg P \leq B_{\text{num}} := \begin{cases} 2d & (b = 2), \\ 4d/b^{r-1} & (b \geq 3), \end{cases}$$

$$(5.2) \quad \deg Q \leq B_{\text{den}} := \begin{cases} 2(1 - 1/2^r)d & (b = 2), \\ 3d/b^{r-1} & (b \geq 3). \end{cases}$$

Proof. Note that, by Lemma 5.1, $\gcd(P, Q) = 1$ implies $\gcd(M^{r-1}P, M^{r-1}Q) = 1$. For $n \in \mathbb{N}$, let $y^{\overline{n}}$ denote the “rising factorial” $\prod_{i=0}^{n-1} M^i y$. In particular, $y^{\overline{0}} = 1$. Multiplying (R) by $Q^{\overline{r}}$ yields

$$(5.3) \quad \sum_{i=0}^r \ell_i P^{\overline{i}} (M^i Q)^{\overline{r-i}} = 0.$$

Since the factor $M^{r-1}Q$ appears in all terms on the left-hand side of (5.3) but the term for $i = r$, and since $M^{r-1}Q$ and $M^{r-1}P$ are coprime, $M^{r-1}Q$ must divide $\ell_r P^{\overline{r-1}}$. Hence, since $\deg \ell_r \leq d$, the degrees of Q and P are related by the inequality

$$(5.4) \quad b^{r-1} \deg Q \leq d + \frac{b^{r-1} - 1}{b - 1} \deg P.$$

Furthermore, for (5.3) to hold, at least one term of index $i < r$ must have a degree greater than or equal to the degree of the term of index r . Simplifying by the common factor P^i yields

$$\deg \left(\ell_r (M^i P)^{\overline{r-i}} \right) \leq \deg \left(\ell_i (M^i Q)^{\overline{r-i}} \right),$$

from which follows

$$\deg P - \deg Q \leq \frac{\deg \ell_i - \deg \ell_r}{(b^r - b^i)/(b-1)} \leq \frac{d}{b^{r-1}}.$$

The above inequality yields an upper bound on $\deg P$,

$$(5.5) \quad \deg P \leq \frac{d}{b^{r-1}} + \deg Q.$$

Substituting it into (5.4), we find

$$b^{r-1} \deg Q \leq d + \frac{b^{r-1} - 1}{b-1} \left(\frac{d}{b^{r-1}} + \deg Q \right) = d \frac{b^r - 1}{(b-1)b^{r-1}} + \frac{b^{r-1} - 1}{b-1} \deg Q.$$

Multiplying by $b-1$ yields

$$(5.6) \quad (b^r - 2b^{r-1} + 1) \deg Q \leq \frac{d}{b^{r-1}} (b^r - 1) \leq db.$$

For the case $b \geq 3$, using the last term, db , in (5.6), and the inequality $b^r \leq 3(b^r - 2b^{r-1} + 1)$ yields (5.2). For the case $b = 2$, specializing the left inequality in (5.6) at $b = 2$ yields precisely (5.2). In both cases, (5.1) is a direct consequence of (5.5) and (5.2). \square

Remark 5.3. The elementary result of Proposition 5.2 is remarkable as it provides a uniform polynomial bound in terms of scalar parameters describing the size of the equation, namely its order r and degree d . This is unreachable in the shift operator case, where the degrees of rational solutions of Riccati equations may be exponential in the bit size of the equation. For example, the Charlier polynomials

$$C_n(x, a) = {}_2F_0 \left(\begin{matrix} -n, -x \\ - \end{matrix} ; -a^{-1} \right)$$

viewed as functions of x with parameters $n \in \mathbb{N}$ and $a > 0$ satisfy the recurrence equation (NIST 2010, §18.22)

$$aC_n(x+1, a) - (a+x)C_n(x, a) + xC_n(x-1, a) + nC_n(x, a) = 0$$

whose degrees of coefficients are bounded by $d = 1$ and whose last coefficient n has bit size $\log n$, while $C_n(x, a)$ has degree n . As a result the associated Riccati equation has solutions $C_n(x+1, a)/C_n(x, a)$ whose numerator and denominator cannot be bounded solely as a function of d . The fact that the Mahler operator drastically increases the degree is sometimes a difficulty, but here it allows us to obtain the bounds in Lemma 5.2.

Remark 5.4. The exponents of b and d in the bounds (5.1) and (5.2) are tight, as we show now. For a given nonzero rational function $u = P/Q$ in its lowest terms

with both P and Q monic, let us consider any operator of order $r \geq 2$

$$\begin{aligned} L &= \left(\sum_{k=0}^{r-1} c_k M^k \right) (QM - P) \\ &= c_{r-1} (M^{r-1} Q) M^r + \sum_{k=1}^{r-1} (c_{k-1} M^{k-1} Q - c_k M^k P) M^k - c_0 P, \end{aligned}$$

with constant coefficients c_0, \dots, c_{r-1} in $\overline{\mathbb{K}}$ satisfying $c_{r-1} \neq 0$. The operator admits $M - u$ as a right-hand factor, hence u is a solution of the Riccati equation according to Lemma 2.1. The leading coefficient of L has degree $b^{r-1} \deg Q$. In case $\deg Q \geq b \deg P$, all other coefficients have lower degree. In case $\deg Q < b \deg P$, the coefficient of M^{r-1} has degree $b^{r-1} \deg P$ and the remaining ones have lower degree. Hence, in both cases the maximum degree of the coefficients of L is equal to $d = b^{r-1} \max(\deg P, \deg Q)$. Upon choosing $b \geq 3$ and $\deg P = \deg Q$, the bounds (5.1) and (5.2) become $B_{\text{num}} = 4 \deg P$ and $B_{\text{den}} = 3 \deg Q$, showing that they overshoot by no more than a constant factor 4.

6. MAHLERIAN VARIANT OF PETKOVŠEK'S ALGORITHM

In this section, we present an algorithm for computing hypergeometric solutions of linear Mahler equations adapted from Petkovšek's algorithm for difference equations in the usual shift operator. In doing so, we generalize to arbitrary order a previous adaptation of Petkovšek's algorithm to Mahler equations of order 2 due to Roques.

Equations are given with polynomial coefficients over a field \mathbb{K} , and we solve them for solutions with coefficients in a field \mathbb{L} satisfying $\overline{\mathbb{K}} \supset \mathbb{L} \supset \mathbb{K}$. As a consequence of §3.2, particularly the definition (3.8) for q_λ , we have a bound

$$(6.1) \quad q_{\mathbb{L}} = \text{lcm}_{\lambda \in \mathbb{L} \cap \Lambda} q_\lambda$$

on the ramification orders of solutions in $\mathbb{K}(x^{1/*})$ of the Riccati equation, so that computing all ramified rational solutions reduces to computing the plain rational solutions of a modified equation. Additionally, the method to be used in §6.2 and §6.3 requires a supplementary ramification in its intermediate calculations: whatever the target ramification order, the working order has to be multiplied by a factor b^{r-1} .

6.1. Petkovšek's classical algorithm. In the case of the linear classical difference equation

$$(6.2) \quad \ell_r(x)y(x+r) + \dots + \ell_0(x)y(x) = \sum_{i=0}^r \ell_i(x)y(x+i) = 0$$

and the corresponding Riccati equation

$$(6.3) \quad \ell_r(x)u(x) \cdots u(x+r-1) + \dots + \ell_1(x)u(x) + \ell_0(x) = \sum_{i=0}^r \ell_i(x) \prod_{j=0}^{i-1} u(x+j) = 0,$$

an algorithm due to Petkovšek (1992) is known to solve (6.3) for all its rational function solutions, or equivalently to find all first-order right-hand factors of (6.2).

The algorithm is based on the concept of a *Gosper–Petkovšek form*: for any rational function $u(x) \in \mathbb{L}(x)_{\neq 0}$, there exist a constant $\zeta \in \mathbb{L}_{\neq 0}$ and monic polynomials $A(x), B(x), C(x)$ in $\mathbb{L}[x]$ satisfying:

- (1) $u(x) = \zeta \frac{C(x+1)}{C(x)} \frac{A(x)}{B(x)}$,
- (2) $A(x)$ and $C(x)$ are coprime,
- (3) $B(x)$ and $C(x+1)$ are coprime,
- (4) $A(x)$ and $B(x+i)$ are coprime for all $i \geq 0$.

Now, any potential nonzero rational solution u of the Riccati equation (6.3) leads to the necessary relation

$$(6.4) \quad \sum_{i=0}^r \ell_i(x) \zeta^i C(x+i) \left(\prod_{j=0}^{i-1} A(x+j) \right) \left(\prod_{j=i}^{r-1} B(x+j) \right) = 0.$$

Here, $A(x)$ appears in all the terms of the sum but the one for $i = 0$. As it is coprime to all forward shifts of $B(x)$ and to $C(x)$, it must divide $\ell_0(x)$. Similarly, $B(x+r-1)$ appears in every term but the one for $i = r$, and so must divide $\ell_r(x)$. This motivates iterating on all pairs of monic divisors of ℓ_0 and ℓ_r . Given monic $A(x) \mid \ell_0(x)$ and $B(x) \mid \ell_r(x-r+1)$, where divisibility is meant in $\mathbb{L}[x]$, the leading coefficient of the left-hand side of (6.4) is independent of the choice of a monic $C(x)$. So this leading coefficient yields an algebraic equation in ζ . Each choice of a solution $\zeta \in \mathbb{L}$ turns (6.4) into an equation that can be solved for polynomial solutions $C(x) \in \mathbb{L}[x]$. The algorithm then gathers and returns all found $(\zeta, A(x), B(x), C(x))$.

Remark that the classical definition of the literature, quantified over all integers $i \in \mathbb{N}$ and reproduced as point (4) above, is stronger than needed: the proof has used the property for $0 \leq i \leq r-1$ only. In the Mahler case, it will prove important to use the weaker constraint to get an algorithm. The reader should also compare point (4) of the present section with its analogues in §6.2 and §6.3.

6.2. Roques’s algorithm for order 2. Inspired by Hendriks’ works for usual difference equations (Hendriks 1998) and for q -difference equations (Hendriks 1997), Roques (2018, §6.2) recently presented an analogue of Petkovšek’s algorithm for Mahler equations of order 2, that is, equations of the form

$$(6.5) \quad \ell_2(x)u(x)u(x^b) + \ell_1(x)u(x) + \ell_0(x) = 0.$$

Roques’s original presentation determines along his algorithm a suitable extension of \mathbb{K} sufficient to obtain all solutions from $\overline{\mathbb{K}}(x^{1/*})$. Here, we present a variant that computes with an input field \mathbb{L} , so as to be consistent with the rest of our text.

After finding a bound q such that all solutions $u \in \mathbb{L}(x^{1/*})$ are in fact in $\mathbb{L}(x^{1/q})$, let us introduce a new indeterminate t for which $x = t^{qb}$, so as to prove, for any given $u \in \mathbb{L}(x^{1/*})$, the existence of a nonzero constant $\zeta \in \mathbb{L}_{\neq 0}$ and monic polynomials $A, B, C \in \mathbb{L}[t]$ satisfying

- (1) $u(x) = \zeta \frac{C(x^{1/q})}{C(x^{1/qb})} \frac{A(x^{1/qb})}{B(x^{1/qb})}$, that is, $u(t^{qb}) = \zeta \frac{C(t^b)}{C(t)} \frac{A(t)}{B(t)}$,
- (2) A and C are coprime,
- (3) B and MC are coprime,
- (4) A and $M^i B$ are coprime for all $i \in \{0, 1\}$,
- (5) C and MC are coprime,

where M acts on $\mathbb{L}[t]$ by substituting t^b for t .

Now, any nonzero ramified rational solution u of the Riccati equation (6.5) leads to the necessary relation

$$(6.6) \quad \ell_2(t^{qb})\zeta^2 C(t^{b^2})A(t)A(t^b) \\ + \ell_1(t^{qb})\zeta C(t^b)A(t)B(t^b) + \ell_0(t^{qb})C(t)B(t)B(t^b) = 0.$$

This time, one finds that $A(t)$ must divide $\ell_0(t^{qb})$, while $B(t^b)$ must divide $\ell_2(t^{qb})$, implying that $B(t)$ must divide $\ell_2(t^q)$.

Roques's approach continues in a way similar in spirit to Petkovšek's, although technical reasons require to exchange the order of steps. Roques indeed finds a linear constraint on the degree of C by putting apart the term with ζ^2 in (6.6). Finding the coefficients of C then amounts to solving a finite linear system. As a side remark, Roques's text strictly speaking confuses valuations and degrees and obscures the calculation of the coefficient ζ . We will correct this in the next section.

6.3. A new algorithm for higher-order Mahler equations. In this section, we deal with Riccati Mahler equations (R) of general order $r \geq 2$. The core of the section describes an algorithm for solving for (plain) rational functions from $\mathbb{L}(x)$.

In order to go beyond order 2, we observe that point (4) in Roques's definition requires a coprimality only for $0 \leq i \leq r - 1 = 1$ (when $r = 2$), whereas point (4) in the shift situation requires a coprimality for $0 \leq i$, although the proof in §6.1 for an equation of order r only uses the cases $0 \leq i \leq r - 1$. This motivates us to introduce the following concept of a bounded Gosper–Petkovšek form.

Definition 6.1. *Given a rational function $u(x) \in \mathbb{L}(x)_{\neq 0}$ and an integer $r \geq 2$, a bounded Gosper–Petkovšek form of order r for $u(x)$ is a tuple $(\zeta, A, B, C) \in \mathbb{L}_{\neq 0} \times \mathbb{L}[t]^3$, with A, B, C monic polynomials in a new indeterminate t , such that:*

- (1) $u(t^{b^{r-1}}) = \zeta \frac{C(t^b)}{C(t)} \frac{A(t^{b^{r-1}})}{B(t)}$,
- (2) $M^{r-1}A$ and C are coprime,
- (3) B and MC are coprime,
- (4) $M^i A$ and B are coprime for all $i \in \{0, \dots, r-1\}$,

where M acts on $\mathbb{L}[t]$ by substituting t^b for t .

By Lemma 5.1, point (4) above can be restated equivalently into: $M^{r-1}A$ and $M^i B$ are coprime for all $i \in \{0, \dots, r-1\}$. Also remark that for a polynomial triple (A, B, C) satisfying our definition for $r = 2$, the polynomial triple (MA, B, C) satisfies Roques's definition (with $q = 1$), at least provided $\gcd(C, MC) = 1$. A constructive proof of the existence of bounded Gosper–Petkovšek forms will be provided in §6.4.

Now, consider any potential nonzero rational solution u of the Riccati equation (R), represented by one of its bounded Gosper–Petkovšek forms. Substituting the Gosper–Petkovšek form for u and $t^{b^{r-1}}$ for x in (R) leads after cancelling denominators to the necessary relation

$$(6.7) \quad \tilde{L}(t, \zeta M) C = 0,$$

where \tilde{L} is the operator in $\mathbb{L}[t]\langle M \rangle$ defined by

$$(6.8) \quad \tilde{L}(t, M) = \sum_{k=0}^r \left(M^{r-1} \ell_k(t) \times \prod_{j=0}^{k-1} M^{r-1+j} A \times \prod_{j=k}^{r-1} M^j B \right) M^k.$$

In this formula, note that beside the polynomials A and B of the ring $\mathbb{L}[t]$, we have noted $\ell_k(t)$ for the result of the substitution of t for x in ℓ_k . The factor $M^{r-1}A$ appears in all the terms of the expansion of (6.7) but the one when $k = 0$,

$$M^{r-1}\ell_0(t) \times C \times \prod_{j=0}^{r-1} M^j B;$$

using points (2) and (4) in Definition 6.1, $M^{r-1}A$ must divide $M^{r-1}\ell_0$, and Lemma 5.1 implies that A must divide ℓ_0 . Similarly, $M^{r-1}B$ appears in all the terms of (6.7) but the one when $k = r$,

$$M^{r-1}\ell_r(t) \times \zeta^r M^r C \times \prod_{j=0}^{r-1} M^{r-1+j} A;$$

thus, $M^{r-1}B$ must divide $M^{r-1}\ell_r$, and Lemma 5.1 implies that B must divide ℓ_r .

To complete the pairs (A, B) into candidate tuples (ζ, A, B, C) delivering rational functions in $\mathbb{L}(t)$, it is appropriate to first determine the possible degrees of C for each possible ζ . This leads us to consider the upper Newton polygon of the operator \tilde{L} and the characteristic polynomial associated to each of its edges. The approach parallels that described in §3.2 to find the possible valuation and leading coefficient λ of a Puiseux series solution of (L) using the lower Newton polygon. For future reference, for a general L , we denote $\xi_j(X)$ the *characteristic polynomial* of the j th edge of the upper Newton polygon of L . Define $Z(L)$ to be the union of the sets of roots in $\overline{\mathbb{K}}$ of the ξ_j . The j th edge will be called ζ -*admissible* (for L) if $\xi_j(\zeta) = 0$.

The previous considerations lead to Algorithm 1, whose general structure is the following:

- a loop over candidates (A, B) is set up at step (B) from the *lower* Newton polygon of the input L ;
- candidates ζ , then degrees for candidates C at step (B)(a), are obtained from the *upper* Newton polygon of the auxiliary operator \tilde{L} ;
- solving for C at step (B)(3)(b)(i) is done by appealing to our algorithm for polynomial solutions of bounded degree in (CDDM 2018, §2.6, Algorithm 5);
- to avoid redundancy in the output, the cleaning step (C) makes sure to return a partition by enforcing that no parametrization is included in another.

For the last item, an algorithm for testing inclusion is as follows: Given two parametrized rational functions, $u^{(1)}$ with parameters $c^{(1)} = (c_1^{(1)}, \dots, c_{s_1}^{(1)})$ and $u^{(2)}$ with parameters $c^{(2)} = (c_1^{(2)}, \dots, c_{s_2}^{(2)})$, inclusion of $u^{(1)}$ in $u^{(2)}$ is only possible if $s_1 \leq s_2$, in which case getting rid of denominators in the equation $u^{(1)} = u^{(2)}$ and equating like powers of x results in a linear system in $c^{(2)}$ linearly parametrized by $c^{(1)}$. Either this system has no nonzero solution, proving noninclusion, or it provides a parametrization proving inclusion. Step (C) now removes redundant element of \mathcal{U} as follows: while \mathcal{U} contains two distinct elements $u^{(1)}$ and $u^{(2)}$ with inclusion of $u^{(1)}$ in $u^{(2)}$, remove $u^{(1)}$ from \mathcal{U} .

We now prove a first part of the correctness of the algorithm: no rational solution is lost.

Input: A Riccati Mahler equation (R) with coefficients $\ell_k(x) \in \mathbb{K}[x]$. Some intermediate field \mathbb{L} , that is, a field satisfying $\bar{\mathbb{K}} \supset \mathbb{L} \supset \mathbb{K}$.

Output: The set of rational functions $u \in \mathbb{L}(x)$ that solve (R).

- (A) Set $\mathcal{U} := \emptyset$.
- (B) For each monic factor $A(t) \in \mathbb{L}[t]$ of $\ell_0(t)$, for each monic factor $B(t) \in \mathbb{L}[t]$ of $\ell_r(t)$ such that $M^i A$ and B are coprime for $0 \leq i < r$:
- (1) compute $\tilde{L}(t, M)$ by (6.8),
 - (2) compute the upper Newton polygon of \tilde{L} , the set $Z(\tilde{L}) \cap \mathbb{L}$ of roots ζ in \mathbb{L} of the associated characteristic polynomials,
 - (3) for each ζ in $Z(\tilde{L}) \cap \mathbb{L}$:
 - (a) compute the maximum Δ_ζ of the integer values of the opposites of the slopes of the ζ -admissible edges (for \tilde{L}),
 - (b) if $\Delta_\zeta \geq 0$:
 - (i) compute a basis $(C_i)_{1 \leq i \leq s}$ of solutions in $\mathbb{L}[t]_{\leq \Delta_\zeta}$ of $\tilde{L}(t, \zeta M) C = 0$,
 - (ii) if $s > 0$:
 - (α) set $C := \sum_{i=1}^s c_i C_i$ for formal parameters c_i ,
 - (β) normalize the rational function

$$\tilde{u}(t) := \zeta \frac{C(t^b)}{C(t)} \frac{A(t^{b^{r-1}})}{B(t)},$$
 which is an element of $\mathbb{L}(c_1, \dots, c_s)(t^{b^{r-1}})$, so as to identify $u(x) \in \mathbb{L}(c_1, \dots, c_s)(x)$ such that $u(t^{b^{r-1}}) = \tilde{u}(t)$,
 - (γ) augment \mathcal{U} with u .
- (C) Remove redundant elements from \mathcal{U} by the method described before Proposition 6.2 and return the resulting \mathcal{U} .
-

Algorithm 1: Rational solutions to a Riccati Mahler equation. Compare with the efficiency improvements in Algorithm 3.

Proposition 6.2. *Algorithm 1 computes the set of rational solutions of the Riccati equation (R) as a union of sets parametrized by finite-dimensional \mathbb{L} -projective spaces.*

Proof. Fix some rational solution $u(x)$ of (R). After Definition 6.1, we have proved the following necessary conditions for (ζ, A, B, C) to be a bounded Gosper–Petkovšek form of order r of u : $\zeta \in Z(\tilde{L})$, A divides $\ell_0(t)$, B divides $\ell_r(t)$, C is a nonzero polynomial solution of (6.7). Observe that the algorithm iterates on all tuples $(\zeta, A, B, \Gamma(c))$ in $\mathbb{L}_{\neq 0} \times \mathbb{L}[t]^2 \times \mathbb{L}(c_1, c_2, \dots)[t]$ such that, for any values of the parameters c_i in \mathbb{L} , the tuple (ζ, A, B, C) where $C = \Gamma(c)$ satisfies: the four necessary conditions, point (4) of the definition, a degree bound on C that is necessarily satisfied by the polynomial kernel of $\tilde{L}(t, \zeta M)$. The algorithm therefore represents a set that is less constrained than the set of bounded Gosper–Petkovšek forms of order r of u ; in particular, it represents them all, and, by the existence of bounded Gosper–Petkovšek forms, the algorithm must find u . Conversely, for any $u(x)$ element of the output \mathcal{U} , obtained by the algorithm from a tuple (ζ, A, B, C) , expanding (6.7) by using (6.8),

then dividing through by the relevant product shows that $u(x)$ solves (R). We have thus proved that Algorithm 1 computes all rational solutions. \square

For the end of the present section, we call a *block* the image of a similarity class of nonzero hypergeometric elements under $y \mapsto My/y$. The following theorem states the correction of Algorithm 1.

Theorem 6.3. *The output \mathcal{U} from Algorithm 1 consists of formal parametrizations of blocks and exactly one for each block. Each such parametrization is a bijection. In equivalent words, Algorithm 1 computes the set of rational solutions of the Riccati equation (R) as a disjoint union of sets bijectively parametrized by finite-dimensional \mathbb{L} -projective spaces.*

Proof. Each element $U(c) = (c_1, \dots, c_s)$ of \mathcal{U} can be interpreted as a parametrization

$$(6.9) \quad \gamma = (\gamma_1 : \dots : \gamma_s) \in \mathbb{P}^{s-1}(\mathbb{L}) \mapsto U(\gamma) = \zeta \frac{MC}{C} \frac{M^{r-1}A}{B} \in \mathbb{L}(x).$$

Fix y_0 , a nonzero solution of $My_0 = \zeta \frac{M^{r-1}A}{B} y_0$. The solutions of $My = U(\gamma)y$ when γ ranges in $(\mathbb{L}^s)_{\neq 0}$ form a vector space \mathfrak{G} that is exactly $C(\mathbb{L}^s)y_0$. Because C is a polynomial in x , $\mathfrak{G}_{\neq 0}$ is included in a similarity class.

By Proposition 6.2, each similarity class $\mathfrak{H}_{\neq 0}$ is such that the finite-dimensional space \mathfrak{H} is covered as a finite union of vector spaces \mathfrak{G} obtained from elements $U \in \mathcal{U}$. Let d denote the dimension of \mathfrak{H} . Assume that no \mathfrak{G} has dimension d . Then, each \mathfrak{G} is in some hyperplane defined by some nonzero linear form ϕ . Fix a coordinate system of \mathfrak{H} and let S be the parametrized curve $t \mapsto (1, t, \dots, t^{d-1})$, whose image is included in \mathfrak{H} . The curve can only meet a given \mathfrak{G} at finitely many intersections, provided by the zeros of the polynomial $\phi(S(t))$. Because the curve has infinitely many points, this contradicts that \mathfrak{H} is covered. So at least one of the \mathfrak{G} is equal to \mathfrak{H} .

By the absence of redundancy enforced at step (C) in the algorithm, no two distinct elements of \mathcal{U} can produce vector spaces \mathfrak{G}_1 and \mathfrak{G}_2 with $\mathfrak{G}_1 \subset \mathfrak{G}_2$. So any vector space \mathfrak{H} of a similarity class $\mathfrak{H}_{\neq 0}$ is obtained exactly once as a \mathfrak{G} , and only such \mathfrak{H} are obtained. The result for blocks then follows.

Because the algorithm construct C in (6.9) as a formal sum $C = \sum_{j=1}^s c_j C_j$ for \mathbb{L} -linearly independent polynomials $C_j \in \mathbb{L}[x]$, the $C_j y_0$ form an \mathbb{L} -basis of $C(\mathbb{L}^s)y_0$. The implied bijection from \mathbb{L}^s to the suitable \mathfrak{H} induces a bijection from $\mathbb{P}^{s-1}(\mathbb{L})$ to the block associated to $\mathfrak{H}_{\neq 0}$. \square

Corollary 6.4. *The set of $\mathbb{L}(x)$ -hypergeometric solutions of the Mahler equation (L) can be described as a union of \mathbb{L} -vector spaces \mathfrak{H} in direct sum parametrized by the output \mathcal{U} as follows: for given $U \in \mathcal{U}$, the vector space \mathfrak{H} is spanned by a basis (y_1, \dots, y_s) where each y_i is any nonzero solution of the equation $My_i = U(e_i)y_i$ for the i th element of the canonical basis.*

Proof. The statement is a direct consequence of the last paragraph in the proof of Theorem 6.3. \square

Remark 6.5 (efficiency of Algorithm 1). The double loop over A and B induces an exponential behavior, making this basic algorithm inefficient. We will discuss several pruning strategies in §6.5. Furthermore, the algorithm considers tuples (ζ, A, B, C) that provide solutions but are not bounded Gosper–Petkovšek forms. These tuples

Input: A Riccati Mahler equation (R) with coefficients $\ell_k(x) \in \mathbb{K}[x]$. Some intermediate field \mathbb{L} , that is, a field satisfying $\bar{\mathbb{K}} \supset \mathbb{L} \supset \mathbb{K}$.

Output: The set of ramified rational functions $u \in \mathbb{L}(x^{1/*})$ that solve (R).

(A) Compute the lower Newton polygon of $L := \sum_{k=0}^r \ell_k(x)M^k$, the associated characteristic polynomials, the set $\Lambda \cap \mathbb{L}$, and the ramification bound $q_{\mathbb{L}}$ by (6.1).

(B) Call Algorithm 1 for computing the rational solutions of $L(x^{q_{\mathbb{L}}}, M)$ in $\mathbb{L}(x)$.

(C) Substitute $x^{1/q_{\mathbb{L}}}$ for x in the obtained solutions and return the resulting set.

Algorithm 2: Ramified rational solutions to a Riccati Mahler equation.

are redundant because the algorithm also produces the same solutions in Gosper–Petkovšek form. From the point of view of efficiency, the problem is not in rejecting tuples that are not bounded forms, but in not computing too many (redundant) candidates in the first place. The necessity to determine the polynomials C late in the algorithm makes it impossible to consider only bounded Gosper–Petkovšek forms, and is the main cause for getting repeated solutions in the course of the algorithm (see also the Remark 6.8 below).

Remark 6.6 (solving large linear systems). When searching for polynomial solutions of the auxiliary equations $\tilde{L}(t, \zeta M)C = 0$, it is crucial for performance to use the fast algorithms introduced in (CDDM 2018). We take the opportunity to clarify a minor point of confusion in that work. In (CDDM 2018, Algorithm 2), we appeal to the linear solving algorithm of Ibarra, Moran, and Hui (1982) for computing the kernel of an $m \times n$ matrix A in $O(m^{\omega-1}n)$ field operations. This algorithm is based on computing an LSP decomposition of A , that is, a decomposition as a product of a lower square matrix L , a semi-upper triangular matrix S , and a permutation matrix P . Strictly speaking, however, this approach is only valid when $m \leq n$ (fewer rows than columns). As our need is for an $m \times n$ matrix A with $m \geq n$ (more rows than columns), a workaround is to apply the LSP decomposition algorithm to the transpose A^T , thus obtaining a PSU decomposition of A , that is, as a permutation matrix P , a semi-lower triangular matrix S , and an upper square matrix U . To solve for the (right) kernel, we observe $\ker A = U^{-1}(\ker S)$. As S reduces to a lower triangular matrix with nonzero diagonal elements when the zero columns are deleted, finding its kernel is immediate. Inverting U takes $O(n^{\omega})$ operations. So $\ker A$ is obtained in $O(m/n)$ square matrix products, hence in complexity $O(mn^{\omega-1})$. Taking the rank r of A into account and refining the analysis yields the complexity $O(mnr^{\omega-2})$ announced in (CDDM 2018, Prop. 2.14).

To compute all *ramified* rational solutions of (R), we now propose Algorithm 2, which is a simple variant of Algorithm 1. By the property of $q_{\mathbb{L}}$ to be a uniform bound on the ramification order of ramified rational solutions, Algorithm 2 is also correct, in the sense that it satisfies an analogue of Theorem 6.3, where rational solutions are replaced by ramified rational solutions. An obvious adaptation of Corollary 6.4 to $\mathbb{L}(x^{1/*})$ -hypergeometric solutions also holds.

6.4. Existence and computation of bounded Gosper–Petkovšek forms.

The following lemma proves the existence of bounded Gosper–Petkovšek forms at all orders, and its proof provides an algorithm for putting a rational function in bounded Gosper–Petkovšek form of a given order.

Lemma 6.7. *Given any two coprime monic polynomials P and Q in $\mathbb{K}[x]$, define a sequence of triples of polynomials given for all $k \in \mathbb{N}_{\neq 0}$ by*

$$(6.10) \quad (A_1, B_1, C_1) = (P, Q, 1)$$

$$(6.11) \quad (A_{k+1}, B_{k+1}, C_{k+1}) = \left(\frac{A_k}{G_k}, \frac{MB_k}{G_k}, MC_k \times (M^0 G_k \cdots M^{k-1} G_k) \right),$$

where $G_k = \gcd(A_k, MB_k)$. Then, for all $k \in \mathbb{N}_{\neq 0}$, $(1, A_k, B_k, C_k)$ is a bounded Gosper–Petkovšek form of order k for the rational function P/Q . Additionally, there exists $k \in \mathbb{N}_{\neq 0}$ satisfying

$$(6.12) \quad A_{k+i} = A_k, \quad B_{k+i} = M^i B_k, \quad C_{k+i} = M^i C_k, \quad \text{for all } i \in \mathbb{N}.$$

Proof. The proof is by induction on k . The case $k = 1$ is immediate by (6.10). For some $k \in \mathbb{N}_{\neq 0}$, assume that $(1, A_k, B_k, C_k)$ is a bounded Gosper–Petkovšek form of order k for the rational function P/Q and that (6.11) holds. This provides the formulas:

$$(6.13) \quad M^{k-1} \frac{P}{Q} = \frac{MC_k}{C_k} \frac{M^{k-1} A_k}{B_k},$$

$$(6.14) \quad \gcd(M^{k-1} A_k, C_k) = 1,$$

$$(6.15) \quad \gcd(B_k, MC_k) = 1,$$

$$(6.16) \quad \gcd(M^i A_k, B_k) = 1 \quad \text{for all } 0 \leq i < k.$$

Then, first observe

$$\frac{MC_{k+1}}{C_{k+1}} = \frac{M^2 C_k}{MC_k} \frac{M^k G_k}{G_k}$$

and apply M to (6.13), so that

$$M^k \frac{P}{Q} = \frac{M^2 C_k}{MC_k} \frac{M^k (A_{k+1} G_k)}{B_{k+1} G_k} = \frac{MC_{k+1}}{C_{k+1}} \frac{M^k A_{k+1}}{B_{k+1}}.$$

Applying Lemma 5.1 with $j = 1$ to (6.14–6.16) implies

$$\gcd(M^{i+1} (A_{k+1} G_k), B_{k+1} G_k) = 1 \quad \text{for all } 0 \leq i < k,$$

$$\gcd(M^k (A_{k+1} G_k), MC_k) = \gcd(B_{k+1} G_k, M^2 C_k) = 1,$$

so that in particular

$$(6.17) \quad \gcd(M^{i+1} A_{k+1}, B_{k+1}) = \gcd(M^{i+1} A_{k+1}, G_k) = \gcd(M^{i+1} G_k, B_{k+1}) = 1,$$

$$(6.18) \quad \gcd(M^k A_{k+1}, MC_k) = \gcd(B_{k+1}, M^2 C_k) = 1.$$

The construction (6.11) ensures $\gcd(M^0 A_{k+1}, B_{k+1}) = 1$, so combining with the left gcd in (6.17) shows (6.16) at $k + 1$. Applying Lemma 5.1 with $j = k - i - 1$ to the middle gcd in (6.17) next yields

$$\gcd(M^k A_{k+1}, M^{k-i-1} G_k) = 1.$$

Considering those gcd for $0 \leq i < k$ as well as the left gcd in (6.18) proves (6.14) at $k + 1$. Combining the right gcd in (6.17) for $0 \leq i < k$ together with the right

gcd in (6.18) proves (6.15) at $k + 1$. We have obtained that $(1, A_{k+1}, B_{k+1}, C_{k+1})$ is a bounded Gosper–Petkovšek form of order $k + 1$ for the rational function P/Q .

Last, in view of (6.11), the existence of k satisfying (6.12) is equivalent to the existence of k such that $G_{k+i} = 1$ for all $i \in \mathbb{N}$. This k exists because $\deg A_k$ cannot decrease indefinitely. \square

Remark 6.8. Bounded Gosper–Petkovšek forms are not unique, even for a given order. An example with $b = 2$ is given for order $r = 3$ by $(\zeta, A, B, C) = (1, x + 1, 1, 1)$ and $(\zeta', A', B', C') = (1, 1, 1, x^4 - 1)$, where the first form is computed by the direct use of the recurrence in Lemma 6.7.

Remark 6.9. Point (5) in Roques’s definition (§6.2) is not enforced in our definition of bounded Gosper–Petkovšek forms. The mere construction of C_{k+1} in (6.11) makes the coprimality between C and MC generally impossible, at least for $r \geq 3$.

6.5. Optimizations. As experiments show, a direct implementation of Algorithm 1 is inefficient because of the large number of pairs of divisors (A, B) it has to consider. In the present section, we describe several ways to mitigate this issue.

6.5.1. Ensuring coprimality. Given a polynomial $f \in \mathbb{L}[t]$, let $\text{irred}(f)$ denote the set of its monic irreducible factors and $\text{val}_p f$ denote the multiplicity of a monic irreducible polynomial p in f . By representing a monic divisor of ℓ_0 by the family $(\alpha_p)_{p \in \text{irred}(\ell_0)}$ of exponents in its monic irreducible factorization $\prod_{p \in \text{irred}(\ell_0)} p^{\alpha_p}$, the set of such divisors can be viewed as the Cartesian product

$$\mathcal{A} = \prod_{p \in \text{irred}(\ell_0)} [0, \text{val}_p \ell_0],$$

where $[a, b]$ denotes the integer interval $\{a, \dots, b\}$. In the same way, the set of monic divisors of ℓ_r is represented by

$$\mathcal{B} = \prod_{q \in \text{irred}(\ell_r)} [0, \text{val}_q \ell_r].$$

So, the main loop (B) in Algorithm 1 can be viewed as parametrized by a pair (α, β) in the product $\mathcal{A} \times \mathcal{B}$. However, we are only interested in pairs (A, B) that satisfy $\gcd(M^s A, B) = 1$ for all $0 \leq s < r$, and the basic algorithm has to test all those gcd in the double loop over A and B .

To explain how to prune $\mathcal{A} \times \mathcal{B}$, let us imagine that ℓ_r has a monic irreducible factor q with multiplicity $b > 0$ and that ℓ_0 has a monic irreducible factor p with multiplicity $a > 0$ such that q is a factor of some $M^s p$, $0 \leq s < r$. In such a situation, we say that the factor p is *forbidden* (in A) by the factor q (of B). When a pair (α, β) of tuples is in $\mathcal{A} \times \mathcal{B}$, the pair (α_p, β_q) of integers lies in the Cartesian product $[0, a] \times [0, b]$. Nonetheless, only the parts $[0, a] \times \{0\}$ and $\{0\} \times [1, b]$ have to be considered, as (α_p, β_q) outside of these parts violate the coprimality condition. The previous considerations are independent of the other coordinates $\alpha_{p'}$ and $\beta_{q'}$, so only a fraction $(a + b + 1)/((a + 1)(b + 1))$ of the whole product $\mathcal{A} \times \mathcal{B}$ is useful. Taking several pairs (p, q) into account, we expect a small fraction of $\mathcal{A} \times \mathcal{B}$ to remain.

We now make this observation algorithmic. Given a monic irreducible factor q of ℓ_r , we denote by $\mathcal{F}(q)$ the set of monic irreducible factors p of ℓ_0 forbidden by q . Let \mathcal{R} be the subset of $\text{irred}(\ell_r)$ consisting of those factors q of ℓ_r such that $\mathcal{F}(q)$ is nonempty. Given a B , a monic irreducible factor q of B restricts the choice of the

useful A if it is in \mathcal{R} , and places no restriction on A if it is not in \mathcal{R} . Next, for a subset $\pi \subset \mathcal{R}$, let $\mathcal{F}(\pi)$ denote $\bigcup_{q \in \pi} \mathcal{F}(q)$. The set of useful pairs (A, B) can now be seen to be the disjoint union over the subsets $\pi \subset \mathcal{R}$ of the Cartesian products

$$\mathcal{C}_\pi = \prod_{p \in \text{irred}(\ell_0)} \mathcal{A}(p, \pi) \times \prod_{q \in \text{irred}(\ell_r)} \mathcal{B}(\pi, q),$$

where

$$\mathcal{A}(p, \pi) = \begin{cases} [0, \text{val}_p \ell_0] & \text{if } p \notin \mathcal{F}(\pi), \\ \{0\} & \text{if } p \in \mathcal{F}(\pi), \end{cases} \quad \mathcal{B}(\pi, q) = \begin{cases} [0, \text{val}_q \ell_r] & \text{if } q \notin \mathcal{R}, \\ [1, \text{val}_q \ell_r] & \text{if } q \in \pi, \\ \{0\} & \text{if } q \in \mathcal{R} \setminus \pi. \end{cases}$$

Rather than precomputing the $\mathcal{A}(p, \pi)$ and $\mathcal{B}(\pi, q)$ explicitly, in our algorithm we discard useless pairs on the fly.

In the following, we call *factored representation* of a polynomial its representation as a power product of monic irreducible factors. An optimization of Algorithm 1 is obtained by just changing the iteration of loop (B) into

- (B) For all $B := \prod_{q \in \text{irred}(\ell_r)} q^{\beta_q}$ such that $0 \leq \beta_q \leq \text{val}_q \ell_r$ for all $q \in \text{irred}(\ell_r)$:
- (u) let F be the union of the $\mathcal{F}(q)$ over all $q \in \text{irred}(\ell_r)$ such that $\beta_q > 0$,
 - (v) for $p \in \text{irred}(\ell_0)$, set a_p to 0 if $p \in F$, to $\text{val}_p \ell_0$ otherwise,
 - (w) for all $A := \prod_{p \in \text{irred}(\ell_0)} p^{\alpha_p}$ such that $0 \leq \alpha_p \leq a_p$ for all $p \in \text{irred}(\ell_0)$,
- (1) compute $\tilde{L}(t, M)$ by (6.8),
- ...

Here, by the definition of $\mathcal{B}(\pi, q)$, those q for which $\beta_q > 0$ are either in π or in $\text{irred}(\ell_r) \setminus \mathcal{R}$. Therefore, the set F computed at step (u) is $\mathcal{F}(\pi)$, because $\mathcal{F}(q) = \emptyset$ if $q \notin \mathcal{R}$. The integer a_p defined at step (v) reflects the definition of $\mathcal{A}(p, \pi)$.

Additionally, the set $\mathcal{F}(q)$ for a given q can be computed efficiently by using the Gräffe operator G defined by $Gf = \text{Res}_y(y^b - x, f(y))$ (cf. CDDM 2018, §3.2, especially Lemma 3.1), as expressed in the following lemma. The point of the approach is to avoid carrying out divisions of high-degree polynomials $M^s p$ by the polynomials q .

Lemma 6.10. *For any monic irreducible factor q of ℓ_r , one has*

$$\mathcal{F}(q) = \text{irred}(\ell_0) \cap \{q, \sqrt{G}q, \dots, \sqrt{G}^{r-1}q\},$$

where $\sqrt{G}f$ denotes the squarefree part of Gf , computed by forcing exponents to 1 in the factored representation of Gf .

Proof. From the definition of G follow the multiplicativity formula $G(fg) = GfGg$ and the degree-preservation formula $\deg Gp = \deg p$. In particular, G preserves divisibility. One has the crucial relations $u \mid M^i G^i u$ and $G^i M^i u = u^{b^i}$ for any monic irreducible polynomial u and $i \in \mathbb{N}$. In terms of \sqrt{G} , the formulas become $u \mid M^i \sqrt{G}^i u$ and $\sqrt{G}^i M^i u = u$ for $i \in \mathbb{N}$.

Fix an integer s such that $0 \leq s < r$. If $q \mid M^s p$, then upon applying \sqrt{G}^s we find $\sqrt{G}^s q \mid \sqrt{G}^s M^s p = p$, so $\sqrt{G}^s q \mid p$. Conversely, if $\sqrt{G}^s q \mid p$, then upon applying M^s we find $q \mid M^s \sqrt{G}^s q = M^s p$, so $q \mid M^s p$. Now, a monic irreducible factor p of ℓ_0 is forbidden by a monic irreducible factor q of ℓ_r if and only if there

exists an integer s such that $0 \leq s < r$ and $q \mid M^s p$, that is, if and only if there exists an integer s such that $0 \leq s < r$ and $p = \sqrt{G}^s q$. \square

6.5.2. Avoiding redundant pairs. Because some of the coprimality conditions defining bounded Gosper–Petkovšek forms cannot be taken into account in Algorithm 1 until parametrizations of solutions C have been obtained, solutions are naturally found with repetitions. Here, we alleviate this by predicting some of the repetitions from divisibility conditions on A and B . By the design of Algorithm 1, a rational solution $u(x)$ will be generated multiple times (before the final step that removes redundancy) if and only if two triples (A, B, C) and (A', B', C') of monic polynomials satisfying

$$(6.19) \quad \frac{MC}{C} \frac{M^{r-1}A}{B} = \frac{MC'}{C'} \frac{M^{r-1}A'}{B'}$$

are considered for the same ζ during the run of the algorithm.

Recall that in Algorithm 1, ramification is taken care of by introducing a new indeterminate t that plays the role of some appropriate root of x . The results described in §6.5.1 for polynomials in $\mathbb{L}[x]$ apply with trivial modifications for polynomials of $\mathbb{L}[t]$, and for the end of §6.5, we continue with the indeterminate t . In particular, we let ℓ_0 and ℓ_r denote $\ell_0(t)$ and $\ell_r(t)$, respectively.

We will study several scenarios in which, if the pair (A, B) is considered in the run of the algorithm and, for a certain ζ , leads to a polynomial C and a rational solution u , then another triple (A', B', C') leads to the same u for the same ζ . We will thus obtain rules that attempt to: minimize the multiplicity of t in A , maximize the multiplicity of t in B , remove factors of the form Mp/p from A , introduce factors of the form Mp/p in B , replace a divisor Mp with the factor p in A , replace the factor p with a divisor Mp in B . In all cases, the polynomial C will be adjusted to compensate for the change. This compensation will operate by multiplications/divisions by elements of $\mathbb{L}[t]$ obtained independently of ζ . This makes the actual value of ζ irrelevant when considering equalities of the form (6.19), which justifies our pruning independently of ζ .

Pruning rules. The following four situations provide instances of (6.19) that will permit us to discard pairs (A, B) . Note that in each case, monic A , B , C , and p induce monic A' , B' , and C' :

- If $B = p\tilde{B}$, define $A' = A$, $B' = Mp\tilde{B}$, $C' = pC$ to get:

$$\frac{MC}{C} \frac{M^{r-1}A}{B} = \frac{MpMC}{MpC} \frac{M^{r-1}A'}{p\tilde{B}} = \frac{MC'}{C'} \frac{M^{r-1}A'}{B'}.$$

- If $A = Mp\tilde{A}$, define $A' = p\tilde{A}$, $B' = B$, $C' = M^{r-1}pC$ to get:

$$\frac{MC}{C} \frac{M^{r-1}A}{B} = \frac{MC}{C} \frac{M^{r-1}(Mp\tilde{A})}{B} = \frac{M(M^{r-1}pC)}{M^{r-1}pC} \frac{M^{r-1}(p\tilde{A})}{B} = \frac{MC'}{C'} \frac{M^{r-1}A'}{B'}.$$

- If $p \mid Mp$, define $A' = A$, $B' = (Mp/p)B$, $C' = pC$ to get:

$$\frac{MC}{C} \frac{M^{r-1}A}{B} = \frac{MpMC}{pC} \frac{M^{r-1}A}{(Mp/p)B} = \frac{MC'}{C'} \frac{M^{r-1}A'}{B'}.$$

- If $p \mid Mp$ and $A = (Mp/p)\tilde{A}$, define $A' = \tilde{A}$, $B' = B$, $C' = M^{r-1}pC$ to get:

$$\frac{MC}{C} \frac{M^{r-1}A}{B} = \frac{MC}{C} \frac{M^{r-1}((Mp/p)\tilde{A})}{B} = \frac{M(M^{r-1}pC)}{M^{r-1}pC} \frac{M^{r-1}A'}{B} = \frac{MC'}{C'} \frac{M^{r-1}A'}{B'}.$$

For any monic divisor A of ℓ_0 , any monic divisor B of ℓ_r , these guarded formulas provide a rule to discard the pair (A, B) if there exists a monic irreducible polynomial p satisfying any of the predicates:

- (P1) B is of the form $p\tilde{B}$ and $Mp\tilde{B}$ divides ℓ_r .
- (P2) A is of the form $Mp\tilde{A}$ and $p\tilde{A}$ divides ℓ_0 .
- (P3) p divides Mp and $(Mp/p)B$ divides ℓ_r .
- (P4) p divides Mp and A is of the form $(Mp/p)\tilde{A}$.

As an example, we legitimize the use of (P4) after assuming the last situation listed above. Because $A' \mid A \mid \ell_0$ and $B' = B \mid \ell_r$, if (A, B, C) is considered in the algorithm for a certain ζ and leads to a solution u , the pair (A', B') is also considered, and the solution u found with (A, B) after the polynomial solving step gets C will also be found with (A', B') and the same ζ after the polynomial solving step gets C' . Using the pair (A', B') instead of (A, B) is better, because it potentially leads to more solutions, including polynomials C' not divisible by $M^{r-1}p$. The use of the other predicates is justified by a similar reasoning.

Furthermore, although this reasoning does not need the irreducibility of p , one easily proves that there exists a monic irreducible p making a rule apply if and only if there exists a monic (not necessarily irreducible) p making the same (generalized) rule apply. So it is algorithmically pertinent to restrict to monic p .

Finally, making $p = t$ in (P3) and (P4) results in slightly more explicit formulations, respectively:

- (P5) $\text{val}_t B \leq \text{val}_t \ell_r - b + 1$.
- (P6) $\text{val}_t A \geq b - 1$.

Note that making $p = t$ in (P1) results in specific cases of (P5), and doing so in (P2) results in specific cases of (P6).

Iteration of the rules. An optimization for factors is obtained from two distinct monic irreducible polynomials p and q satisfying $Mp = pq$: considering the final result of a repeated use of predicates makes it possible to fix certain multiplicities before entering the loops over A and B . If for $k \geq 1$, $q^k B$ divides ℓ_r , an iterated use of (P3) shows that all of $B, \dots, q^{k-1}B$ can be skipped, and we get the predicate:

- (P7) $Mp = pq$ for $p \neq q$ and $\text{val}_q B < \text{val}_q \ell_r$.

Similarly, if for $k \geq 1$, A is of the form $q^k \tilde{A}$ and divides ℓ_0 , an iterated use of (P4) shows that all of $q^k \tilde{A}, \dots, q \tilde{A}$ can be skipped, and we get the predicate:

- (P8) $Mp = pq$ for $p \neq q$ and $\text{val}_q A > 0$.

Note that $p \neq q$ requires $p \neq t$ and $q \neq t$.

Implementation of the discarding rule. The optimization of loop (B) that we introduced in §6.5.1 is further refined in Algorithm 3: before (u) in the loop over B , we insert a step (t) to take (P1) and (P3) into account, and before (1) in the loop over A , we insert a step (0) to take (P2) and (P4) into account. Notwithstanding, the specializations (P5) to (P8) are taken care of at step (Z), by restricting the exponents used in A and B to ranges in which the rules (P1) to (P4) cannot apply.

Because (P3) tests those polynomials Mp/p that divide ℓ_r , a precomputation at step (k) determines those quotients. For any monic irreducible factor q of any such Mp/p , $\sqrt{G}q$ is equal to p , thus restricting the search for p to $\text{irred}(G\ell_r)$. A similar discussion applies to (P4) and step (l).

6.5.3. Avoiding redundant computations of Newton polygons. Step (B)(1) of Algorithm 1 computes the characteristic polynomials $\xi_j(X)$ associated to \tilde{L} for every pair (A, B) of divisors of ℓ_0 and ℓ_r . It turns out that the roots of the $\xi_j(X)$ do not depend on (A, B) . In this section, the *degree* of a finite Puiseux series is defined as the maximal (rational) exponent appearing with a nonzero coefficient. Also, given an operator P , we write $Z(P)$ for the union of the sets of roots in $\overline{\mathbb{K}}$ of the characteristic polynomials associated with all edges of the upper Newton polygon of P .

Lemma 6.11. *The set $Z(\tilde{L})$ is equal to $Z(L)$ for every pair (A, B) of polynomials in $\mathbb{L}[t]$. Moreover, the set $\Delta(\tilde{L})$ of degrees of finite Puiseux series solutions of $\tilde{L}y = 0$ is deduced from the set $\Delta(L)$ of degrees of finite Puiseux series solution of $Ly = 0$ by an affine transform:*

$$(6.20) \quad \Delta(\tilde{L}) = \text{sh}(\Delta(L)) \quad \text{where} \quad \text{sh}(\delta) = -\frac{b^{r-1} \deg A - \deg B}{b-1} + b^{r-1}\delta.$$

Proof. The upper Newton polygon of L , respectively \tilde{L} , is the upper part of the convex hull of the points (b^k, d_k) , $0 \leq k \leq r$, respectively of the points (b^k, D_k) , $0 \leq k \leq r$, with the relation

$$D_k = b^{r-1}d_k + b^{r-1}\frac{b^k - 1}{b-1} \deg A + \frac{b^r - b^k}{b-1} \deg B = b^{r-1}d_k + \alpha b^k + \beta$$

that results directly from (6.8) after setting

$$\alpha = \frac{b^{r-1} \deg A - \deg B}{b-1} \quad \text{and} \quad \beta = \frac{b^r \deg B - b^{r-1} \deg A}{b-1}.$$

As the coefficient b^{r-1} is positive, the upper convex hull of the points (b^k, d_k) , $0 \leq k \leq r$, is mapped onto the upper convex hull of the points (b^k, D_k) , $0 \leq k \leq r$. As a consequence, the degrees of the finite Puiseux series solutions of L and \tilde{L} are related by (6.20), since they are the opposite of the slopes of the edges (CDDM 2018, Lemma 2.5). Moreover, as the polynomials A and B are monic, the coefficients of the monomials $x^{d_k}M^k$ in L and $x^{D_k}M^k$ in \tilde{L} are the same for $0 \leq k \leq r$. So the characteristic polynomials are the same for the edges of both upper Newton polygons. \square

A first consequence of Lemma 6.11 is that we can change the structure of Algorithm 1: instead of recomputing $Z(\tilde{L})$ at step (2) for each new (A, B) , a single computation of $Z(L)$ is done before entering loop (B), and step (3) becomes a loop over $\zeta \in Z(L)$. In the latter loop, we still need to compute a bound on the degree of a solution C for each pair (A, B) , but it is deduced from the lemma by considering the largest nonnegative integer in $\Delta(\tilde{L})$ as obtained from $\Delta(L)$ by the parametrization (6.20).

Input: A Riccati Mahler equation (R) with coefficients $\ell_k(x) \in \mathbb{K}[x]$. Some intermediate field \mathbb{L} , that is, a field satisfying $\mathbb{K} \supset \mathbb{L} \supset \mathbb{K}$.

Output: The set of rational functions $u \in \mathbb{L}(x)$ that solve (R).

- (W) Compute the upper Newton polygon of $L := \sum_{k=0}^r \ell_k(x)M^k$, the associated characteristic polynomials, the set $Z \cap \mathbb{L}$ where $Z = Z(L)$, and for each $\zeta \in Z \cap \mathbb{L}$, the set of indices j making the j th edge ζ -admissible.
 - (X) From now on, let ℓ_k denote $\ell_k(t)$.
 - (Y) Compute factored representations of relevant polynomials and related data:
 - (j) for $q \in \text{irred}(\ell_r)$, set $\mathcal{F}(q) := \text{irred}(\ell_0) \cap \{q, \sqrt{G}q, \dots, \sqrt{G}^{r-1}q\}$,
 - (k) set \mathcal{D}_r to the set of those $f \in \{Mp/p : p \in \text{irred}(G\ell_r), p \neq t\} \cap \mathbb{L}[t]$ that divide ℓ_r , [(P3), not (P5) or (P7)]
 - (l) set \mathcal{D}_0 to the set of those $f \in \{Mp/p : p \in \text{irred}(G\ell_0), p \neq t\} \cap \mathbb{L}[t]$ that divide ℓ_0 . [(P4), not (P6) or (P8)]
 - (Z) Refine bounds for the loops below:
 - (j) if $t \mid \ell_r$, set $\check{b}_t := \max(0, \text{val}_t \ell_r - b + 2)$, [(P5)]
 - (k) if $t \mid \ell_0$, set $\check{a}_t := b - 2$, [(P6)]
 - (l) for $q \in \text{irred}(\ell_r)$ with $q \neq t$, set \check{b}_q to $\text{val}_q \ell_r$ if $Mp = pq$ for $p = \sqrt{G}q$, to 0 otherwise, [(P7)]
 - (m) for $q \in \text{irred}(\ell_0)$ with $q \neq t$, set \check{a}_q to 0 if $Mp = pq$ for $p = \sqrt{G}q$, to $\text{val}_q \ell_0$ otherwise. [(P8)]
 - (A) Set $\mathcal{U} := \emptyset$.
 - (B) For all $B := \prod_{q \in \text{irred}(\ell_r)} q^{\beta_q}$ such that $\check{b}_q \leq \beta_q \leq \text{val}_q \ell_r$ for all $q \in \text{irred}(\ell_r)$:
 - (t) continue to the next B if either of the following conditions holds:
 - (α) $Mp(B/p) \mid \ell_r$ for some $p \in \text{irred}(B)$ with $p \neq t$, [(P1), not (P5)]
 - (β) $fB \mid \ell_r$ for some $f \in \mathcal{D}_r$, [(P3), not (P5) or (P7)]
 - (u) let F be the union of the $\mathcal{F}(q)$ over all $q \in \text{irred}(\ell_r)$ such that $\beta_q > 0$,
 - (v) for $p \in \text{irred}(\ell_0)$, set a_p to 0 if $p \in F$, to $\text{val}_p \ell_0$ otherwise,
 - (w) for all $A := \prod_{p \in \text{irred}(\ell_0)} p^{\alpha_p}$ such that $0 \leq \alpha_p \leq \min(\check{a}_p, a_p)$ for all $p \in \text{irred}(\ell_0)$,
 - (0) continue to the next A if either of the following conditions holds:
 - (α) for $p \in \text{irred}(\ell_0)$ with $p \neq t$, Mp divides A and $p(A/Mp)$ divides ℓ_0 , [(P2), not (P6)]
 - (β) $fA \mid \ell_0$ for some $f \in \mathcal{D}_0$, [(P4), not (P6) or (P8)]
 - (1) compute $\check{L}(t, M)$ by (6.8),
 - (2) for each ζ in $Z \cap \mathbb{L}$:
 - (a) compute the maximum Δ_ζ of the integer values taken by $\text{sh}(\delta)$ in (6.20) when δ ranges over the opposites of the slopes of the ζ -admissible edges of L ,
 - (b) proceed as in step (B)(3)(b) of Algorithm 1.
 - (C) Finish as in step (C) of Algorithm 1.
-

Algorithm 3: Improved variant of Algorithm 1.

7. ALTERNATIVE ALGORITHM BY HERMITE-PADÉ APPROXIMATION

We develop another approach to finding all ramified rational function solutions of the Riccati Mahler equation (R). In comparison to the adaption of Petkovšek's algorithm in §6, which focuses on fixing the possible singularities of a solution of (R), this second approach will proceed in a guess-and-check manner, refining calculations on approximate solutions of the linear equation (L) until a whole set of candidate solutions is proven to be exact solutions. This approach will be embodied in our Algorithm 4, which guarantees to get all solutions.

Our algorithm will search for the set $\mathfrak{A}_{\overline{\mathbb{K}}(x^{1/s})}$ of solutions with coefficients in $\overline{\mathbb{K}}$, but for the first part of the theory, we more generally consider an intermediate field $\overline{\mathbb{K}} \supset \mathbb{L} \supset \mathbb{K}$, as in §6. We compute Hermite–Padé approximants for auxiliary problems, before recombining them to obtain structured approximants. These auxiliary problems show no logarithm and ramification, making it possible to focus on formal power series calculations. Each auxiliary problem solves a modified Riccati equation for its unramified rational solutions whose series expansions have nonnegative valuation and their leading coefficient is equal to 1.

Let us focus on the space of formal power series solutions of the linear equation (L) and consider an \mathbb{L} -basis (z_1, \dots, z_t) , from which we deduce solutions u of (R) in $\mathbb{L}[[x]]$ by the parametrization

$$u = \frac{M(a_1 z_1 + \dots + a_t z_t)}{a_1 z_1 + \dots + a_t z_t}, \quad a = (a_1 : \dots : a_t) \in \mathbb{P}(\mathbb{L}^t).$$

We need to determine those a for which the series u lies in $\mathbb{L}(x)$. Suppose that u is indeed some P/Q in $\mathbb{L}(x)$, with coprime P and Q . After chasing denominators and using linearity, we obtain the equation

$$(7.1) \quad (-a_1 P) z_1 + \dots + (-a_t P) z_t + (a_1 Q) M z_1 + \dots + (a_t Q) M z_t = 0.$$

This linear relation with polynomial coefficients from $\mathbb{L}[x]$ between the series $z_1, \dots, z_t, M z_1, \dots, M z_t$ from $\mathbb{L}[[x]]$ is a special case of a linear relation

$$(7.2) \quad P_1 z_1 + \dots + P_t z_t + Q_1 M z_1 + \dots + Q_t M z_t = 0,$$

for polynomial coefficients P_i and Q_i from $\mathbb{L}[x]$. A second level of relaxation is obtained, for any $\sigma \in \mathbb{N}$, by Hermite–Padé approximants $(P_1, \dots, P_t, Q_1, \dots, Q_t)$ in $\mathbb{L}[x]^{2t}$, to order σ , that is, approximate linear relations of the form

$$(7.3) \quad P_1 z_1 + \dots + P_t z_t + Q_1 M z_1 + \dots + Q_t M z_t = O(x^\sigma).$$

Recombining exact relations (7.2) into *structured* relations of the form (7.1) reduces to solving a polynomial system in a_1, \dots, a_t (see §7.2), so that we can in principle go back from (7.2) to (7.1). Unfortunately, given bounds on the degrees of the P_i and Q_i , it is not clear how to find an accuracy σ such that (7.3) implies (7.2). Instead, Algorithm 4 computes the approximate relations (7.3) compatible with our degree bounds on P and Q and attempts to reconstruct hypergeometric solutions starting from these relations. This proceeds by guessing candidate relations for increasing σ and rejecting wrong ones until we get no false solutions. We will show (Theorem 7.16) that this process eventually yields all hypergeometric solutions, and nothing but hypergeometric solutions.

The same idea can be adapted to find the rational solutions of a linear Mahler equation. We further comment on this in §7.6. The latter, as well as §7.5 on the rank of relations module, are only incidental to the flow of the text.

Remark 7.1. The problem of determining those a making u rational simplifies a lot when $t = 1$, as it indeed reduces to the question whether $M z_1 / z_1$ is rational. Using Hermite–Padé approximation together with the bounds derived in §5 give an efficient and simple procedure in that context.

7.1. Approximate syzygies. In order to introduce the suitable modules where the tuples $(P_1, \dots, P_t, Q_1, \dots, Q_t)$ of the previous paragraph can be found, we recall the definition of a *syzygy module* associated with a tuple of series. Given an integer $m \in \mathbb{N}$ and a tuple $f = (f_1, \dots, f_m) \in \mathbb{L}[[x]]^m$, an element $w = (p_1, \dots, p_m) \in \mathbb{L}[x]^m$ is a

syzygy of f when $w \cdot f^T = 0$. The set of syzygies is an $\mathbb{L}[x]$ -submodule of $\mathbb{L}[x]^m$, denoted $\text{Syz}(f_1, \dots, f_m)$. Given an integer $\sigma \in \mathbb{N}$, we also consider the module $\mathcal{S}^{[\sigma]}$ of *approximate syzygies* of f , defined as the set of $w \in \mathbb{L}[x]^m$ for which there exists a series $q \in \mathbb{L}[[x]]$ satisfying $(w, q) \cdot (f, x^\sigma)^T = 0$. Note that this is in general larger than the projection to its first m components of $\text{Syz}(f_1, \dots, f_m, x^\sigma)$, as the latter would restrict the implied q to polynomials. By definition, an approximate syzygy $w = (p_1, \dots, p_m) \in \mathcal{S}^{[\sigma]}$ satisfies $w \cdot f^T = O(x^\sigma)$. Hence, the sequence of modules is nonincreasing: for any $\tau \leq \sigma$, $\mathcal{S}^{[\tau]} \supset \mathcal{S}^{[\sigma]}$. It is well known (Derksen 1994) that $\mathcal{S}^{[\sigma]}$ is a free $\mathbb{L}[x]$ -module of rank m , and that this module admits a basis (w_1, \dots, w_m) such that for each i , the m -tuple $w_i = (p_1, \dots, p_m)$ satisfies $\deg p_j \leq \deg p_i > \deg p_{j'}$ whenever $j \leq i < j'$. Such bases were later called *minimal bases* in the literature. They are also well known to be Gröbner bases of the approximate syzygy module for a term-over-position (TOP) ordering (Neiger 2016). Fast algorithms for their calculation have been provided, most notably in (Beckermann and Labahn 1994); see also (Beckermann and Labahn 2000).

To link the previous definitions with the notation of the introduction of §7, we set $m := 2t$ and, for $1 \leq i \leq t$, $f_i := z_i$ and $f_{i+t} := Mz_i$. If the Riccati equation has nontrivial solutions, for a large enough σ , we expect the existence of an approximate syzygy of the form

$$(7.4) \quad w_0 := (-a_1P, \dots, -a_tP, a_1Q, \dots, a_tQ)$$

that is in fact an exact syzygy satisfying $\deg P \leq B_{\text{num}}$ and $\deg Q \leq B_{\text{den}}$ for the bounds introduced in Proposition 5.2. Let us consider a minimal basis $(w_1^{[\sigma]}, \dots, w_{2t}^{[\sigma]})$ of $\mathcal{S}^{[\sigma]}$. Any $w \in \mathcal{S}^{[\sigma]}$ reduces to zero by reduction by the minimal basis viewed as a Gröbner basis for the TOP order. Thus, there are polynomials R_i such that

$$w = R_1w_1^{[\sigma]} + \dots + R_{2t}w_{2t}^{[\sigma]}$$

with $\deg(R_iw_i^{[\sigma]}) \leq \deg w$ for $1 \leq i \leq 2t$: for the quotient R_i to be nonzero, we need $\deg w_i^{[\sigma]} \leq \deg w$. In particular, for any $d \in \mathbb{N}$, any w of degree at most d reduces to zero by the elements of the minimal basis themselves of degree at most d . With our goal to be able to generate structured syzygies w_0 of degree bounded as in (7.4), what just precedes justifies retaining only those elements of the minimal basis whose degrees are at most

$$(7.5) \quad B_\infty := \max(B_{\text{num}}, B_{\text{den}}).$$

The approximate syzygy module truncated to degree B_∞ is the \mathbb{L} -vector space

$$(7.6) \quad \mathcal{T}^{[\sigma]} := \mathcal{S}^{[\sigma]} \cap \mathbb{L}[x]_{\leq B_\infty}^{2t} = \sum_{\deg w_i^{[\sigma]} \leq B_\infty} \mathbb{L}[x]_{\leq B_\infty - \deg w_i^{[\sigma]}} w_i^{[\sigma]}.$$

Without loss of generality we can assume that $v := \text{val } z_1 \leq \text{val } z_i$ for $1 \leq i \leq t$. Thus, for $\sigma \geq v$, $(x^{\sigma-v}, 0, \dots, 0)$ is in $\mathcal{S}^{[\sigma]}$ but not in $\mathcal{S}^{[\sigma+1]}$. The sequence of the modules $\mathcal{S}^{[\sigma]}$ ultimately decreases, while $\mathcal{S}^{[\sigma]}$ retains its rank. Let $\mathcal{S}^{[\infty]}$ denote the limit $\bigcap_{\sigma \in \mathbb{N}} \mathcal{S}^{[\sigma]}$, which is nothing but $\text{Syz}(f_1, \dots, f_m)$, because a series is 0 if and only if it is $O(x^\sigma)$ for all $\sigma \in \mathbb{N}$. The sequence of the $\mathcal{T}^{[\sigma]}$ is nonincreasing, and, as a subspace of a fixed finite-dimensional vector space, is ultimately constant. Its limit $\mathcal{T}^{[\infty]}$ is $\text{Syz}(f_1, \dots, f_m) \cap \mathbb{L}[x]_{\leq B_\infty}^{2t}$, and the module it spans, $\mathbb{L}[x]\mathcal{T}^{[\infty]}$, is an $\mathbb{L}[x]$ -submodule of $\text{Syz}(f_1, \dots, f_m)$. We have just proven:

Lemma 7.2. *For each $\sigma \in \mathbb{N}$, the $\mathbb{L}[x]$ -module $\mathbb{L}[x]\mathcal{F}^{[\sigma]}$ is generated by those row vectors $w_i^{[\sigma]} \in \mathbb{L}^{2t}$ of a minimal basis of $\mathcal{S}^{[\sigma]}$ that satisfy $\deg w_i^{[\sigma]} \leq B_\infty$. The sequence of these modules is nonincreasing and ultimately constant with limit the submodule $\mathbb{L}[x]\mathcal{F}^{[\infty]}$ of $\text{Syz}(f_1, \dots, f_m)$. A similar property holds for the sequence of their ranks, which we write*

$$(7.7) \quad 2t \geq \rho^{[\sigma]} := \text{rank } \mathbb{L}[x]\mathcal{F}^{[\sigma]} \geq \rho^{[\infty]} := \text{rank } \mathbb{L}[x]\mathcal{F}^{[\infty]}.$$

To continue with an algorithmic description in the next sections, we write $W^{[\sigma]}$ for the $\rho^{[\sigma]} \times (2t)$ -matrix whose rows are those $w_i^{[\sigma]}$ with degree at most B_∞ , for $1 \leq i \leq 2t$.

Remark 7.3. Given a tuple $f = (f_1, \dots, f_m) \in \mathbb{L}[[x]]^m$, some precision σ , and a minimal basis $w = (w_1, \dots, w_m)$ of the corresponding $\mathbb{L}[x]$ -module $\mathcal{S}^{[\sigma]}$ of approximate syzygies, it is immediate, by reduction, that the basis w is also a minimal basis of its $\mathbb{L}'[x]$ -module of approximate syzygies at precision σ for any superfield \mathbb{L}' of \mathbb{L} .

7.2. From unstructured to structured syzygies. Given an operator $L \in \mathbb{K}[x]\langle M \rangle$ of order r with a space of formal Laurent series solutions assumed of \mathbb{L} -dimension $t > 0$, a suitable composition on the right by a power of x ensures that all its Laurent series solutions are in fact formal power series and that the minimal valuation among solutions is 0. In this situation, we fix a basis $z := (z_1, \dots, z_t)$ of the space of series solutions. For any $\sigma \in \mathbb{N}$, the $\mathbb{L}[x]$ -module $\mathbb{L}[x]\mathcal{F}^{[\sigma]}$ of unstructured (approximate) syzygies may contain structured syzygies. We consider the set

$$(7.8) \quad \mathcal{A}^{[\sigma]} = \{0\} \cup \{a \in (\mathbb{L}^t)_{\neq 0} \mid \exists P \in \mathbb{L}[x]_{\neq 0}, \exists Q \in \mathbb{L}[x]_{\neq 0}, \\ (-Pa, Qa) \in \mathbb{L}[x]\mathcal{F}^{[\sigma]}\},$$

which is easily seen to be an \mathbb{L} -cone. As a consequence of Lemma 7.2, the sequence of the cones $\mathcal{A}^{[\sigma]}$ is nonincreasing and ultimately constant, with limit a cone $\mathcal{A}^{[\infty]}$ that we proceed to describe.

The following lemma makes properties of $\mathcal{A}^{[\infty]}$ explicit, as consequences of Theorem 2.8.

Lemma 7.4. *Let s_1, s_2, \dots denote the dimensions of the nontrivial similarity classes $(\mathfrak{H}_1)_{\neq 0}, (\mathfrak{H}_2)_{\neq 0}, \dots$ of the hypergeometric series solutions of (\mathbf{L}) . There exist full-rank matrices H_1, H_2, \dots such that:*

- the spaces $\mathbb{L}^{s_i} H_i$ are in direct sum in \mathbb{L}^t ;
- the cone $\mathcal{A}^{[\infty]}$ is the union of the \mathbb{L} -vector spaces $\mathbb{L}^{s_i} H_i$.

Proof. Consider an $\mathbb{L}(x)$ -similarity class of hypergeometric series solutions $\mathfrak{H}_{\neq 0}$ of (\mathbf{L}) , described by a basis $h := (h_1, \dots, h_s)$. Expressing this family in the basis z provides us with a matrix $H \in \mathbb{L}^{s \times t}$ such that $h^T = Hz^T$. Each $c \in (\mathbb{L}^s)_{\neq 0}$ yields a rational-function solution $M(ch^T)/(ch^T)$ of (\mathbf{R}) , also equal to $M(cHz^T)/(cHz^T)$. After writing P/Q for this rational function, we get that $(-PcH, QcH)$ is a structured approximate syzygy to any order σ , so that cH is in the cone $\mathcal{A}^{[\infty]}$. This cone therefore contains $\mathbb{L}^s H$. Iterating over the nontrivial similarity classes $(\mathfrak{H}_1)_{\neq 0}, (\mathfrak{H}_2)_{\neq 0}, \dots$ of hypergeometric series solutions, we therefore get dimensions s_1, s_2, \dots and matrices H_1, H_2, \dots such that each $\mathbb{L}^{s_i} H_i$ is contained in $\mathcal{A}^{[\infty]}$. By point 1 of Theorem 2.8 applied to $D = \mathfrak{D}_{\mathbb{L}}$ and $F = \mathbb{L}((x))$, the \mathfrak{H}_i are in direct sum in $\mathfrak{D}_{\mathbb{L}}$, inducing, by the map $a \mapsto az^T$, that the $\mathbb{L}^{s_i} H_i$ are in direct sum in the \mathbb{L} -space

generated by $\mathcal{A}^{[\infty]}$. Conversely, any nonzero $a \in \mathcal{A}^{[\infty]}$ yields a nonzero rational solution P/Q of the Riccati Mahler equation, thus, by point 2 of Theorem 2.8 must come from some $(\mathfrak{H}_i)_{i \neq 0}$. So, the cone $\mathcal{A}^{[\infty]}$ consists solely of the union of the \mathbb{L} -vector spaces $\mathbb{L}^{s_i} H_i$. \square

The goal of the end of the present §7.2 is to study properties of the sequence of the cones $\mathcal{A}^{[\sigma]}$ and related objects, which will justify our Algorithm 4 in §7.4: our algorithm will work by increasing σ until obtaining $\mathcal{A}^{[\sigma]} = \mathcal{A}^{[\infty]}$. What follows contains sufficient conditions detecting that σ is too small to have $\mathcal{A}^{[\sigma]} = \mathcal{A}^{[\infty]}$.

In view of (7.7), if some rank $\rho^{[\sigma]}$ is 0, then $\rho^{[\infty]} = 0$, meaning there are no nontrivial relations at all, and our algorithm will terminate in this case. So, we continue the analysis by assuming $\rho^{[\infty]} > 0$. At the other extreme, if the rank $\rho^{[\infty]}$ were $2t$, we could produce $2t$ nonzero polynomials p_i such that $p_i z_i = 0$ for $1 \leq i \leq 2t$, which cannot be. Thus, we have $0 < \rho^{[\infty]} < 2t$. Consequently, we also have $\rho^{[\sigma]} < 2t$ for large enough σ , and our algorithm ensures to work in this case.

Example 7.5. The operator whose origin will be described in Example 8.7 will serve as a working example throughout §7.2. It has order $r = 4$, degree $d = 258$, and its space of formal power series solutions in $\mathbb{Q}[[x]]$ has dimension $t = 4$. For any $\sigma \leq 120$, the computation of the matrix $W^{[\sigma]}$ leads to a rank $\rho^{[\sigma]} = 8 = 2t$, proving that such σ are too low to provide candidates.

From now on, we assume $1 \leq \rho^{[\sigma]} \leq 2t - 1$. The following lemma is a variant of Cramer's rule.

Lemma 7.6. *Let M be a matrix of size $(n + 1) \times n$ and rank n , with coefficients in a field F . Then, the left kernel of M has dimension 1 over F and a nonzero kernel element is $K := (\Delta_1, \dots, (-1)^{i+1} \Delta_i, \dots, (-1)^n \Delta_{n+1})$, where Δ_i denotes the determinant of the square submatrix obtained by removing the i th row.*

Proof. Augmenting M by any column C of it on its left yields a matrix with determinant 0. Expanding the determinant with respect to the first column shows that K satisfies $KC = 0$. Combining all columns C yields $KM = 0$. Because $\text{rk } M = n$, the minors Δ_i cannot all be 0 simultaneously, which implies $K \neq 0$. \square

The nonzero elements a of the cone $\mathcal{A}^{[\sigma]}$ defined by are characterized by the presence of a structured row $(-Pa, Qa)$ in.

In terms of the basis of the module $\mathbb{L}[x] \mathcal{T}^{[\sigma]}$ provided by the rows of $W^{[\sigma]}$, the cone $\mathcal{A}^{[\sigma]}$ defined by (7.8) admits the equivalent definition

$$(7.9) \quad \mathcal{A}^{[\sigma]} = \{0\} \cup \{a \in (\mathbb{L}^t)_{\neq 0} \mid \exists P \in \mathbb{L}[x]_{\neq 0}, \exists Q \in \mathbb{L}[x]_{\neq 0}, \exists \Lambda \in \mathbb{L}[x]^{\rho^{[\sigma]}}, \\ (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, P, -Q)W_+^{[\sigma]} = 0\},$$

where $W_+^{[\sigma]}$ is the matrix $W^{[\sigma]}$ augmented at its bottom with the two-row matrix

$$(7.10) \quad \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} a_1 & \dots & a_t & 0 & \dots & 0 \\ 0 & \dots & 0 & a_1 & \dots & a_t \end{pmatrix}.$$

Note that this alternative definition is in fact independent of the specific choice of the basis in W . We introduce the cone $\mathcal{V}^{[\sigma]}$ defined analogously by the looser constraint $(\Lambda, P, -Q) \neq 0$ in place of $P \neq 0$ and $Q \neq 0$: this is the cone of values of a that make $W_+^{[\sigma]}$ have a nontrivial left kernel. Note the inclusion $\mathcal{A}^{[\sigma]} \subset \mathcal{V}^{[\sigma]}$. The following lemma describes a case of equality.

Lemma 7.7. *For σ large enough, meaning $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$, assume $1 \leq \rho^{[\sigma]} \leq 2t - 1$. Then:*

- (1) *if, for some nonzero a , the row $K = (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, -P, Q) \neq 0$ satisfies $KW_+^{[\sigma]} = 0$, then neither P nor Q can be zero;*
- (2) *$\mathcal{A}^{[\sigma]} = \mathcal{V}^{[\sigma]} = \mathcal{A}^{[\infty]}$.*

Proof. Taking a as in the first point implies $Paz^T = QaMz^T + \Lambda W^{[\sigma]}(z, Mz)^T$, which reduces to $Paz^T = QaMz^T$ because $W^{[\sigma]}(z, Mz)^T = 0$ by the hypothesis on σ . Now, if P is zero, then $Qa = 0$ because the entries of Mz are independent over \mathbb{L} , forcing $Q = 0$ because $a \neq 0$. Similarly, $Q = 0$ implies $P = 0$. So, if P or Q is zero, then both must be zero, which contradicts that $W^{[\sigma]}$ has independent rows over $\mathbb{L}(x)$, thus proving $\mathcal{V}^{[\sigma]} \subset \mathcal{A}^{[\sigma]}$, and therefore $\mathcal{A}^{[\sigma]} = \mathcal{V}^{[\sigma]}$. Moreover, if two integers σ_1 and σ_2 are such that $\mathcal{F}^{[\sigma_1]} = \mathcal{F}^{[\sigma_2]}$, then $\mathcal{A}^{[\sigma_1]} = \mathcal{A}^{[\sigma_2]}$ by the definition (7.8). In particular, for all σ satisfying $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$, all the cones $\mathcal{A}^{[\sigma]}$ are equal and equal to $\mathcal{A}^{[\infty]}$. \square

Two cases need different analyses: (i) $\rho^{[\sigma]} = 2t - 1$; (ii) $\rho^{[\sigma]} \leq 2t - 2$.

7.2.1. Case $\rho^{[\sigma]} = 2t - 1$. In the first case, the $(\rho^{[\sigma]} + 2) \times (\rho^{[\sigma]} + 1)$ -matrix $W_+^{[\sigma]}$ has a nontrivial left kernel for any value of a : the cone $\mathcal{V}^{[\sigma]}$ is the whole space \mathbb{L}^t . Additionally, the $\rho^{[\sigma]} \times (\rho^{[\sigma]} + 1)$ -matrix $W^{[\sigma]}$ has full rank over $\mathbb{L}(x)$, and thus has a 1-dimensional right kernel. Applying Lemma 7.6 to the transposed matrix provides a polynomial generator $K^T = (\Delta_1, \dots, (-1)^{i+1}\Delta_i, \dots, -\Delta_{2t})^T$ of this right kernel, given by minors Δ_i obtained by removing columns. As ranks are unchanged by extending the base field to $\mathbb{L}((x))$, $W^{[\sigma]}$ also has a 1-dimensional right kernel over $\mathbb{L}((x))$, and the same K^T is a generator of this $\mathbb{L}((x))$ -vector space.

Lemma 7.8. *For σ large enough, meaning $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$, assume $\rho^{[\sigma]} = 2t - 1$. Then, $\mathcal{A}^{[\sigma]} = \mathcal{V}^{[\sigma]} = \mathbb{L}^t$, and those solutions u to the Riccati equation (R) that are simultaneously quotients Mw/w where $w = \sum_{i=1}^t a_i z_i$ for some $a \in (\mathbb{L}^t)_{\neq 0}$ and rational solutions in $\mathbb{L}(x)$ are given by the parametrization*

$$(7.11) \quad (a_1 : \dots : a_t) \mapsto u = \frac{\sum_{i=1}^t a_i (-1)^{i+t+1} \Delta_{i+t}}{\sum_{i=1}^t a_i (-1)^{i+1} \Delta_i}, \quad \text{with } a \in \mathbb{P}^{t-1}(\mathbb{L}).$$

Proof. By the assumption on σ , $(z_1, \dots, z_t, Mz_1, \dots, Mz_t)^T$ must be in the kernel of $W^{[\sigma]}$, and thus must be a multiple of K^T by a series $z_0 \in \mathbb{L}((x))_{\neq 0}$. In particular, $z_i = (-1)^{i+1} \Delta_i z_0$ for $1 \leq i \leq t$, so all the z_i are \mathbb{L} -similar. For any $a \in (\mathbb{L}^t)_{\neq 0}$, the nonzero series $w = \sum_{i=1}^t a_i z_i$ is a solution to (L) and an immediate calculation shows that the quotient Mw/w is a rational solution from $\mathbb{L}(x)$ given by the formula in (7.11). That is, w is hypergeometric and, by point 2 of Theorem 2.8, $u := Mw/w$ is a solution to (R). This shows that (7.11) parametrizes all solutions u that are quotients Mw/w and rational. Any such solution, given by $a \neq 0$, provides nonzero P and Q satisfying $Mw/w = P/Q$. The nullity of $(P, -Q) \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} (z, Mz)^T$ implies that $(P, -Q) \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ is a linear combination of the rows of $W_+^{[\sigma]}$ with rational function coefficients, that is, there exists $\Lambda \in \mathbb{L}(x)^{\rho^{[\sigma]}}$ satisfying $(\Lambda, P, -Q)W_+^{[\sigma]} = 0$. After clearing denominators, this shows $a \in \mathcal{A}^{[\sigma]}$. So we proved the equality $\mathcal{A}^{[\sigma]} = \mathbb{L}^t$, from which derives the equality $\mathcal{V}^{[\sigma]} = \mathbb{L}^t$ because $\mathcal{A}^{[\sigma]} \subset \mathcal{V}^{[\sigma]} \subset \mathbb{L}^t$. \square

The algorithm does not decide at first if $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$. Instead, the formula (7.11) of Lemma 7.8 introduces a parametrized candidate for rational solutions. If this parametrized rational function is later verified to be an actual parametrized solution, in other words if the algorithm reaches a σ large enough to ensure

$$u = \frac{\sum_i a_i z_{i+t}}{\sum_i a_i z_i},$$

then the similarity class of hypergeometric solutions of the z_i is proved to have \mathbb{L} -dimension t .

Example 7.9 (continuing from Example 7.5). For each σ between 121 and 123, the computation of the matrix $W^{[\sigma]}$ leads to a rank $\rho^{[\sigma]} = 7 = 2t - 1$. In each case, the rational candidate (7.11) has a numerator of degree 81 in x and a denominator of degree 83, both involving linear parameters a_1, \dots, a_4 . It will turn out, however, that this candidate need not parametrize only solutions (see Example 7.13).

7.2.2. Case $1 \leq \rho^{[\sigma]} \leq 2t - 2$. In the second case, the matrix $W_+^{[\sigma]}$ is of size $(\rho^{[\sigma]} + 2) \times (2t)$ with $\rho^{[\sigma]} + 2 \leq 2t$. This matrix is full rank over $\mathbb{L}(x, a_1, \dots, a_t)$ and we want to find all choices of $a = (a_1, \dots, a_t)$ in \mathbb{L}^t that make the rank drop. They are described by cancelling all maximal minors. These minors are polynomials in a and x that are homogeneous of degree 2 in the a_i . We gather their coefficients with respect to x into a system $\Sigma^{[\sigma]}$ of polynomials of $\mathbb{L}[a]$. So, the cone $\mathcal{V}^{[\sigma]}$ of all a that make $W_+^{[\sigma]}$ rank deficient is the variety of common zeros in \mathbb{L}^t of $\Sigma^{[\sigma]}$. The following lemma makes the description of the equality case $\mathcal{A}^{[\sigma]} = \mathcal{V}^{[\sigma]}$ in Lemma 7.7 more explicit.

Lemma 7.10. *For σ large enough, meaning $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$, assume $1 \leq \rho^{[\sigma]} \leq 2t - 2$. Then, for any nonzero a of \mathbb{L}^t that makes $W_+^{[\sigma]}$ rank deficient, the rank over $\mathbb{L}(x)$ of $W_+^{[\sigma]}$ is $\rho^{[\sigma]} + 1$.*

Proof. Because $\text{rk } W^{[\sigma]} = \rho^{[\sigma]}$, the rank $\text{rk } W_+^{[\sigma]}$ is either $\rho^{[\sigma]}$ or $\rho^{[\sigma]} + 1$. In the former case, the first $\rho^{[\sigma]}$ rows of $W_+^{[\sigma]}$ generate its image $\mathbb{L}(x)^{\rho^{[\sigma]}+2} W_+^{[\sigma]}$, and there exists $K = (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, -P, 0) \in \mathbb{L}[x]^{\rho^{[\sigma]}+2}$ with $P \neq 0$ satisfying $KW_+^{[\sigma]} = 0$. Consequently, $Pa z^T = (\Lambda_1 w_1^{[\sigma]} + \dots + \Lambda_{\rho^{[\sigma]}} w_{\rho^{[\sigma]}}^{[\sigma]})(z, Mz)^T$, which is 0 because $W^{[\sigma]}(z, Mz)^T = 0$ by the hypothesis on σ . It follows that $az^T = 0$, which contradicts that z is an \mathbb{L} -basis. So the rank must be $\rho^{[\sigma]} + 1$. \square

By Lemma 7.4, the limit cone $\mathcal{A}^{[\infty]}$ is a union of linear spaces. The cone $\mathcal{V}^{[\sigma]}$ approximates $\mathcal{A}^{[\sigma]}$ by containing it and is equal to it for sufficiently large σ . Our algorithm therefore computes a primary decomposition of the radical of the ideal generated by $\Sigma^{[\sigma]}$, then tests whether each prime ideal thus obtained describes a linear space.

Algorithms for primary decomposition over $\overline{\mathbb{K}}$ exist (Gianni, Trager, and Zacharias 1988; Decker, Greuel, and Pfister 1999). They return primary ideals represented by (minimal reduced) Gröbner bases, and we will prove by Theorem 7.15 that testing linearity of the irreducible varieties of $\mathcal{V}^{[\sigma]}$ amounts to checking that all returned ideals are generated by linear forms.

So, on the one hand, if any element of any of the (minimal reduced) Gröbner bases is nonlinear, then our algorithm restarts with a larger σ .

On the other hand, if all Gröbner bases define linear spaces, our algorithm introduces a more explicit parametrization of candidates $u = P/Q$. To this end, take each Gröbner basis in turn, solve the system it represents for a to determine parameters (g_1, \dots, g_v) and some matrix S of size $v \times t$ such that $a = gS$. The following lemma shows the resulting form of candidate solutions P/Q , parametrized by g .

Lemma 7.11. *Assume $1 \leq \rho^{[\sigma]} \leq 2t - 2$ and let $g \in \mathbb{L}^v$ satisfy $gS \in (\mathcal{V}^{[\sigma]})_{\neq 0}$. Then, the left kernel of $W_+^{[\sigma]}$ is generated over $\mathbb{L}(x)$ by a vector $K = (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, -P, Q)$ where P and Q are nonzero and depend linearly on g , and the Λ_i are quadratic in g .*

Proof. The matrix $W_+^{[\sigma]}$ has size $(\rho^{[\sigma]} + 2) \times (2t)$ and by Lemma 7.10 it has $\mathbb{L}(x)$ -rank $\rho^{[\sigma]} + 1$. As a consequence, we can select $\rho^{[\sigma]} + 1$ linearly independent columns to obtain a matrix $\Omega(g)$ of rank $\rho^{[\sigma]} + 1$, made of an upper $\rho^{[\sigma]} \times (\rho^{[\sigma]} + 1)$ block extracted from $W_+^{[\sigma]}$ and independent of g , and of two rows depending linearly on g . Lemma 7.6 provides a nonzero $K = (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, -P, Q)$ satisfying $K\Omega(g) = 0$, with nonzero P and Q by Lemma 7.7. Since the columns of $\Omega(g)$ generate those of $W_+^{[\sigma]}$, the product $KW_+^{[\sigma]}$ is zero as well. The definition of K by minors and the degree structure of $W_+^{[\sigma]}$ with respect to a provide the degrees in g of its entries. \square

Example 7.12 (continuing from Example 7.9). For $\sigma = 124$, computing the matrix $W^{[\sigma]}$ leads to a rank $\rho^{[\sigma]} = 6 = 2t - 2$. Augmenting $W^{[\sigma]}$ by stacking the two-row matrix (7.10) and taking the coefficients with respect to x of the single minor to be considered yields a system $\Sigma^{[\sigma]}$ of 30 polynomials of degree 2 in a_1, \dots, a_4 . The corresponding irredundant prime decomposition is $\mathfrak{J}_1 \cap \mathfrak{J}_2 \cap \mathfrak{J}_3$ for $\mathfrak{J}_1 = \langle a_1 - a_3, a_2 - a_4 \rangle$, $\mathfrak{J}_2 = \langle a_1 - a_2, a_1 + a_3, a_1 + a_4 \rangle$, $\mathfrak{J}_3 = \langle a_1 - a_4, a_1 + a_2, a_1 + a_3 \rangle$. The associated linear subspaces of \mathbb{L}^4 , predicted to be in direct sum by Lemma 7.4, have respective dimensions 2, 1, and 1. In the present case, they add up to $t = 4$. The ideals \mathfrak{J}_2 and \mathfrak{J}_3 yield isolated candidates P/Q , corresponding to projective dimension 0, while the first ideal yields a candidate parametrized by a projective line. Focusing on \mathfrak{J}_1 , we solve it as $(a_1, a_2, a_3, a_4) = (g_1, g_2) \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$. The specialized augmented matrix becomes

$$W_+^{[\sigma]} = \begin{pmatrix} x & -1 & x & -1 & 0 & 0 & 0 & 0 \\ p_{2,1}^{(36)} & p_{2,2}^{(37)} & p_{2,3}^{(0)} & p_{2,4}^{(37)} & p_{2,5}^{(0)} & p_{2,6}^{(1)} & p_{2,7}^{(0)} & p_{2,8}^{(1)} \\ 0 & 0 & -1 & 0 & x^2 & 0 & 1 & x \\ 0 & 0 & 0 & -1 & 0 & x^2 & x & 1 \\ -1 & 0 & 0 & 0 & 1 & x & x^2 & 0 \\ 0 & -1 & 0 & 0 & x & 1 & 0 & x^2 \\ g_1 & g_2 & g_1 & g_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g_1 & g_2 & g_1 & g_2 \end{pmatrix},$$

where the polynomials $p_{2,j}^{(d_j)}$ in the second row have degrees d_j for $1 \leq j \leq 8$. This matrix has 8 rows and we observe that the first 7 columns are linearly independent over $\mathbb{L}(x, g_1, g_2)$, leading to the 8×7 matrix $\Omega(g)$. The determinants of the submatrices of $\Omega(g)$ obtained by deleting the j th row for $1 \leq j \leq 8$ are all of the form

$\Delta_j = q\delta_j$ with q a degree 36 polynomial and

$$\begin{aligned} \delta_1 &= -(g_1^2 - g_2^2)x, & \delta_2 &= 0, \\ \delta_3 = \delta_5 &= -(g_1 - g_2x + g_1x^2)(g_1 + g_2x), & \delta_4 = \delta_6 &= (g_2 - g_1x + g_2x^2)(g_1 + g_2x), \\ & & \delta_7 &= -(g_1 + g_2x^3), & \delta_8 &= -(1 + x^2 + x^4)(g_1 + g_2x). \end{aligned}$$

The left kernel is the $\mathbb{L}(x)$ -line generated by $K := (\Delta_1, \dots, \Delta_7, -\Delta_8)$. This provides a rational candidate

$$(7.12) \quad \frac{P}{Q} = \frac{\Delta_7}{\Delta_8} = \frac{g_1 + g_2x^3}{g_1 + g_2x} \frac{1}{1 + x^2 + x^4}.$$

7.2.3. Validation of the candidate u . Whether obtained when $\rho^{[\sigma]} = 2t - 1$ (§7.2.1) or $\rho^{[\sigma]} \leq 2t - 2$ (§7.2.2), a candidate $u = P/Q$ has its numerator and denominator parametrized linearly by some $g = (g_1, \dots, g_v)$, potentially after setting $g = a$ and $v = t$ (if in the former case).

Suppose a candidate permitted to have false solutions and true solutions for different values of g . For sure, enlarging σ would ultimately reject the false solutions. Besides, trying to restrict the parametrizing set of g to the correct ones by substituting P/Q for u in (R) and identifying coefficients to 0 would lead to a polynomial system of degree r in g , which we would not know how to solve efficiently.

So, verifying that a sufficiently large σ has been used amounts to verify for each parametrized candidate that P/Q is truly a solution of (R) for *all* choices of g . As a first criterion, any fixed parametrized candidate in normal form P/Q can thus be rejected if either $\deg P > B_{\text{num}}$ or $\deg Q > B_{\text{den}}$, for the bounds defined by (5.1) and (5.2). Otherwise, after substituting P/Q for u in the left-hand side of (R), either we obtain 0 and P/Q is a valid solution, or σ has to be increased.

Example 7.13 (continuing from Example 7.12). Gathering the results from the previous examples, we see that the rank $\rho^{[\sigma]}$ of the approximate syzygy module $\mathcal{F}^{[\sigma]}$ defined by (7.6) decreases by levels when σ increases. Accordingly, the cone $\mathcal{A}^{[\sigma]}$ of parameters of candidate solutions decreases as well, until it reaches the limit $\mathcal{A}^{[\infty]}$ that corresponds to true solutions only. Let us describe this in more detail:

- $\rho^{[\sigma]} = 2t = 8$ for $\sigma \leq 120$. We already explained that these σ are too small.
- $\rho^{[\sigma]} = 2t - 1 = 7$ for $121 \leq \sigma \leq 123$. As the degrees in x of each candidate are $81 > 344/9 = B_{\text{num}} \simeq 38.2$ for the numerator and $83 > B_{\text{den}} = 86/3 \simeq 28.7$ for the denominator, we reject those candidates.
- $\rho^{[\sigma]} = 6 \leq 2t - 2$ for $124 \leq \sigma \leq 126$. For $\sigma = 124$ and the parametrized candidate (7.12), substituting into the Riccati equation proves that (7.12) defines a family of true rational solutions parametrized by $(g_1 : g_2) \in \mathbb{P}^1(\mathbb{Q})$. Similarly, the ideals \mathfrak{J}_2 and \mathfrak{J}_3 of Example 7.12 each yields one solution. As all candidates are solutions, $\mathcal{A}^{[\sigma]} = \mathcal{A}^{[\infty]}$ for this σ , and therefore for all $\sigma \geq 124$.
- $\rho^{[\sigma]} = 5 \leq 2t - 2$ for $127 \leq \sigma$. As we will see in §7.5, each similarity class of hypergeometric solutions of dimension s (over $\mathbb{L} = \mathbb{Q}$) contributes $2s - 1$ linearly independent linear relations (of degree at most $B_\infty = B_{\text{num}} \simeq 38.2$), and in presence of several classes, the $2s - 1$ add up to a lower bound on the rank $\rho^{[\infty]}$ of the submodule $\mathbb{L}[x]\mathcal{F}^{[\infty]}$ of the module of relations on hypergeometric solutions. In the present example, this lower bound on $\rho^{[\infty]}$ is $(2 \times 2 - 1) + (2 \times 1 - 1) + (2 \times 1 - 1) = 5$. A computation for $\sigma = 127$

finds $\rho^{[\sigma]} = 5$, proving that $\rho^{[\sigma]} = \rho^{[\infty]} = 5$, as well as $\mathcal{T}^{[\sigma]} = \mathcal{T}^{[\infty]}$, for all $\sigma \geq 127$.

The reader might be surprised to note that the equality $\mathcal{A}^{[\sigma]} = \mathcal{A}^{[\infty]}$ is satisfied before $\mathcal{T}^{[\sigma]} = \mathcal{T}^{[\infty]}$ occurs: indeed, the definition of $\rho^{[\sigma]}$ as the rank of $\mathbb{L}[x]\mathcal{T}^{[\sigma]}$ and the inequality $\rho^{[124]} = 6 > \rho^{[127]} = 5 = \rho^{[\infty]}$ imply that $\mathcal{T}^{[124]}$ is strictly larger than $\mathcal{T}^{[\infty]}$, whereas $\mathcal{A}^{[124]} = \mathcal{A}^{[\infty]}$.

7.3. Ramified rational solutions. As a consequence of Theorem 3.14, all the ramified rational solutions of (R) with a given leading coefficient λ are images by $y \mapsto My/y$ of the \mathbb{L} -vector space of solutions y of (L) in $(\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/q_\lambda}))$ for q_λ given by (3.8). Furthermore, by the same theorem, the set of $\lambda \in \mathbb{L}$ such that the image $\mathfrak{R}_{\mathbb{L}, \lambda}$ is nonempty is contained in the finite set $\mathbb{L} \cap \Lambda'$, which is computable. Let us fix such a λ . Then, we first change the operator $L = L(x, M)$ into $L(x, \lambda M)$ to reduce the search in $(\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/q_\lambda}))$ to a search in $\mathbb{L}((x^{1/q_\lambda}))$, and we next apply the transformation explained in (CDDM 2018, §2.7, particularly Lemma 2.21) to reduce the search in $\mathbb{L}((x^{1/q_\lambda}))$ to a search in $\mathbb{L}[[x]]$. To make the operator resulting from those combined transformations more explicit: consider the rightmost λ -admissible edge of the lower Newton polygon of L whose denominator is coprime with b ; express its slope as $-p_\lambda/q_\lambda$, which is possible by the definition of q_λ ; and denote by c_λ its intercept on the ordinate axis. By defining

$$(7.13) \quad L_\lambda(x, M) = x^{-q_\lambda c_\lambda} L(x^{q_\lambda}, \lambda M) x^{p_\lambda},$$

where the factor $x^{-q_\lambda c_\lambda}$ simply ensures that the coefficients of L_λ are coprime polynomials, we obtain a linear Mahler operator whose solutions $z \in \mathbb{L}[[x]]$ parametrize the solutions y of L in $(\ln x)^{\log_b \lambda} \mathbb{L}((x^{1/q_\lambda}))$ by $y(x) = (\ln x)^{\log_b \lambda} x^{p_\lambda/q_\lambda} z(x^{1/q_\lambda})$. This parametrization is bijective, owing to (CDDM 2018, Prop. 2.19). To obtain all the ramified rational solutions u of (R), it is therefore sufficient to apply the method of §7.2 to each L_λ and to obtain $u = My/y$ for y defined from z as above.

Example 7.14. To show an example of calculation of L_λ , let us consider the radix $b \geq 2$ and the linear Mahler operator $L = \lambda_0 x^\omega - M + M^r$, with λ_0 an algebraic number, ω a positive integer coprime to $b - 1$, and r an integer larger than 1. The characteristic polynomial of the leftmost edge of the Newton polygon has a single root, λ_0 , and this edge is the only λ_0 -admissible edge. The change described by (7.13) results in the linear operator $L_{\lambda_0} = \lambda_0(1 - M + \lambda_0^{r-1} x^{(b^r-b)\omega} M^r)$. The power series solutions of L_{λ_0} are multiples of

$$(7.14) \quad z = 1 - \lambda_0^{r-1} x^{(b^r-b)\omega} - \lambda_0^{r-1} x^{(b^r-b)b\omega} + \dots,$$

from which we deduce a line of solutions of L in \mathfrak{D} generated by

$$y = (\ln x)^{\log_b \lambda_0} x^{\omega/(b-1)} \left(1 - \lambda_0^{r-1} x^{(b^r-b)\omega/(b-1)} + \dots \right).$$

7.4. Algorithm by sieving candidates. By composing the results developed in the previous subsections, we derive Algorithm 4, in which the body of the loop over the approximation order σ has been isolated as Algorithm 5. The termination and correction of the algorithm will be proved in Theorem 7.16, after having proved a structural property of the ideals it manipulates in Theorem 7.15. Before this, we comment on choices and algorithms used at a few specific steps:

- For each λ , the algorithm computes a basis of power series solutions truncated to an order σ_0 that is sufficient to distinguish the basis elements. This

order is determined at step (C)(c). To explain the calculation, remind (CDDM 2018, Prop. 2.6 and Lemma 2.21) that the computation of a basis of solutions for the equation $L_\lambda z = 0$ depends on the rational numbers

$$(7.15) \quad \nu_\lambda = q_\lambda \nu - p_\lambda, \quad \mu_\lambda = q_\lambda (\mu - c_\lambda),$$

where $-\nu$ and μ are respectively the slope and intercept of the *leftmost* edge of the Newton polygon of L , and, as was explained before (7.13), $-p_\lambda/q_\lambda$ and c_λ are respectively the slope and intercept of the *rightmost* 1-admissible edge of the Newton polygon of L_λ . The order σ_0 is then defined to be $\lfloor \nu_\lambda \rfloor + 1$.

- Solving for truncated solutions z_i at step (C)(d) is done by solving a linear system of size $(\lfloor \nu_\lambda \rfloor + 1) \times (\lfloor \mu_\lambda \rfloor + 1)$ as per (CDDM 2018, §2.6, Algorithm 5).
- The initial value of σ chosen at step (C)(h) uses all data already available at this point. Next, choosing a geometric sequence to grow σ , as opposed to, for example, an arithmetic sequence, is justified by the superlinear complexity with respect to σ of the calculations in the loop body. The golden ratio used to increase σ is rather arbitrary, but reduces a majority of timings for the set of equations tested.
- Prolonging the truncated solutions z_i at step (C)(h)(1) is done by (CDDM 2018, §2.3, Algorithm 3).
- Computing the minimal basis at step (C)(h)(2) can be done by either the algorithm in (Derksen 1994) or the algorithm in (Beckermann and Labahn 1994). In both cases, the theoretical description admits algebraic coefficients in \mathbb{K} .
- At step (C)(h)(6)(γ), the algorithm uses the classical notion of an irredundant primary decomposition of an ideal, which applies in particular to radical ideals, in which case the primary ideals are even prime (Zariski and Samuel 1958, p. 209, especially Theorem 5); we will speak of an irredundant prime decomposition in such a situation. The computation of such a prime decomposition for radical ideals in $\mathbb{K}[a_1, \dots, a_t]$ can be done by a variant of the algorithm in (Gianni, Trager, and Zacharias 1988); see also (Becker and Weispfenning 1993, Algorithms RADICAL, p. 394, and PRIMDEC, p. 396).

The proof of Theorem 7.15 bases on the Nullstellensatz. Consequently, for the rest of §7.4 we specialize \mathbb{L} to $\overline{\mathbb{K}}$ in the theory developed in §7.1 and §7.2.

Theorem 7.15. *Consider an operator $L \in \overline{\mathbb{K}}[x]\langle M \rangle$ and a basis $z := (z_1, \dots, z_t)$ of formal power series solutions in $\overline{\mathbb{K}}[[x]]$. Consider as well the family of modules $\overline{\mathbb{K}}[x]\mathcal{F}^{[\sigma]}$ of ranks $\rho^{[\sigma]}$ and the family of cones $\mathcal{A}^{[\sigma]}$ defined in §7.1 and §7.2. Assume that there exists σ satisfying the hypothesis $1 \leq \rho^{[\sigma]} \leq 2t - 2$ of §7.2.2 and large enough to have the relation $\mathcal{F}^{[\sigma]} = \mathcal{F}^{[\infty]}$. The system $\Sigma := \Sigma^{[\sigma]} \subset \overline{\mathbb{K}}[a] := \overline{\mathbb{K}}[a_1, \dots, a_t]$ defined in §7.2.2 generates an ideal $\langle \Sigma \rangle$ whose radical $\sqrt{\langle \Sigma \rangle}$ is a finite intersection of prime ideals, each given by linear polynomials with coefficients in $\overline{\mathbb{K}}$.*

Proof. Lemma 7.4 proves the existence of a finite family of ideals $\mathfrak{q}_i \subset \overline{\mathbb{K}}[a]$ generated by $\overline{\mathbb{K}}$ -linear polynomials, and satisfying

$$(7.16) \quad \mathcal{A}^{[\infty]} = \bigcup_{i \in I} V(\mathfrak{q}_i),$$

Input: A Riccati Mahler equation (R) with coefficients in $\mathbb{K}[x]$.

Output: The set of ramified rational functions $u \in \overline{\mathbb{K}}(x^{1/*})$ that solve (R).

- (A) Compute the lower Newton polygon \mathcal{N} of $L := \sum_{k=0}^r \ell_k(x)M^k$, then the set Λ by (3.7).
 - (B) Determine the leftmost edge of \mathcal{N} and compute its slope $-\nu$ and intercept μ .
 - (C) For each λ in Λ ,
 - (a) compute the ramification order q_λ as defined by (3.8),
 - (b) determine the rightmost λ -admissible edge of \mathcal{N} and compute its slope $-p_\lambda/q_\lambda$ and intercept c_λ ,
 - (c) compute the operator L_λ by (7.13), the rationals ν_λ and μ_λ given by (7.15) in terms of μ and ν , and the integer $\sigma_0 := \lfloor \nu_\lambda \rfloor + 1$,
 - (d) compute a basis (z_1, \dots, z_t) of solutions in $\mathbb{K}[\lambda][[x]]$ to the equation $L_\lambda z = 0$, truncated to order $O(x^{\sigma_0})$,
 - (e) if $t = 0$, then continue to the next λ ,
 - (f) compute bounds $B_{\text{num}}, B_{\text{den}}, B_\infty$ by (5.1), (5.2), and (7.5) applied to L_λ ,
 - (g) initialize the solution set by $R_\lambda := \emptyset$,
 - (h) for $\sigma := \left(\frac{1+\sqrt{5}}{2}\right)^k \sigma_0$ given by successive $k = 0, 1, 2, \dots$, execute Algorithm 5 in the current context,
 - (i) for each u in R_λ , change $u(x)$ into $x^{(b-1)p_\lambda/q_\lambda} u(x^{1/q_\lambda})$.
 - (D) Return the union of the sets R_λ over $\lambda \in \Lambda$.
-

Algorithm 4: Ramified rational solutions to a Riccati Mahler equation by Hermite–Padé approximants and prime decompositions of radical ideals.

where $V(\cdot)$ denotes the variety in $\overline{\mathbb{K}}^t$ of an ideal. Every $\overline{\mathbb{K}}$ -subspace \mathfrak{H} of $\overline{\mathbb{K}}((x^{1/*}))$ associated to a $\overline{\mathbb{K}}(x^{1/*})$ -similarity class $\mathfrak{H}_{\neq 0}$ of hypergeometric Puiseux series solutions of L admits a finite basis $h = (h_1, \dots, h_s)$ with elements in $\overline{\mathbb{K}}((x^{1/*}))$ that are $\overline{\mathbb{K}}$ -linear combinations of the series z_i . With the notation of the beginning of §7.2, there is a matrix $H \in \overline{\mathbb{K}}^{s \times t}$ such that $h^T = Hz^T$.

Each $V(\mathfrak{q}_i)$ is in bijection with such an \mathfrak{H} by the $\overline{\mathbb{K}}$ -linear map $a \mapsto az^T$, and more specifically $V(\mathfrak{q}_i) = \overline{\mathbb{K}}^s H$. Consequently, \mathfrak{q}_i is generated by a system of $t - s$ linear independent polynomials with coefficients in $\overline{\mathbb{K}}$.

We now have the successive equalities

$$V\left(\sqrt{\langle \Sigma \rangle}\right) = V(\langle \Sigma \rangle) = \mathcal{V}^{[\sigma]} = \mathcal{A}^{[\infty]} = \bigcup_{i \in I} V(\mathfrak{q}_i) = V\left(\bigcap_{i \in I} \mathfrak{q}_i\right),$$

where the first equality is by (Cox, Little, and O’Shea 2015, Theorem 7(ii), p. 183), the second is by definition, the third results by Lemma 7.7 from the assumption $\mathcal{S}^{[\sigma]} = \mathcal{S}^{[\infty]}$, the fourth is (7.16), and the fifth is by (Cox, Little, and O’Shea 2015, Theorem 15, p. 196). Retaining the equality between the first and last terms, then passing to ideals, Hilbert’s Nullstellensatz (Cox, Little, and O’Shea 2015, Theorem 2,

In the context of execution of step (C)(h) in Algorithm 4.

- (1) extend the basis elements z_1, \dots, z_t to solutions truncated to order $O(x^\sigma)$ by unrolling recurrences,
 - (2) compute a minimal basis of the module of approximate syzygies of $f = (z_1, \dots, z_t, Mz_1, \dots, Mz_t)^T$ to order $O(x^\sigma)$,
 - (3) build a matrix W of dimension $\rho \times 2t$ by extracting from the minimal basis the (independent) rows of degree at most B_∞ ,
 - (4) if $\rho = 0$, continue to the next λ ,
 - (5) if $\rho = 2t - 1$:
 - (α) for $i = 1, \dots, 2t$, compute the minor Δ_i obtained after removing the i th column from W ,
 - (β) define $\mathcal{C} := \{u\}$ for the candidate u provided by (7.11), with parameters a_i replaced with g_i ,
 - (6) if $1 \leq \rho \leq 2t - 2$:
 - (α) let W_+ be the $(\rho + 2) \times 2t$ matrix obtained by appending $\begin{pmatrix} a_1 & \dots & a_t & 0 & \dots & 0 \\ 0 & \dots & 0 & a_1 & \dots & a_t \end{pmatrix}$ below W ,
 - (β) let $\langle \Sigma \rangle$ be the ideal of $\overline{\mathbb{K}}[a_1, \dots, a_t]$ generated by the set $\Sigma \subset \mathbb{K}[\lambda][a_1, \dots, a_t]$ of the coefficients with respect to x of the $\binom{2t}{\rho+2}$ minors of order $\rho + 2$ of W_+ ,
 - (γ) compute an irredundant prime decomposition $\bigcap_{j=1}^s \mathfrak{p}_j$ of $\sqrt{\langle \Sigma \rangle}$, given for each j by a Gröbner basis $(p_{j,1}, \dots, p_{j,m(j)})$ of \mathfrak{p}_j ,
 - (δ) initialize the candidate set by $\mathcal{C} := \emptyset$,
 - (ε) for j from 1 to s do
 - [i] if for some k , the polynomial $p_{j,k}$ is nonlinear, continue to the next σ ,
 - [ii] solve the linear system $\{p_{j,k} = 0\}$ for the unknowns a_1, \dots, a_t , so as to get a parametrization $(a_1, \dots, a_t) = (g_1, \dots, g_v)S$ for some full rank matrix S , with $0 \leq v \leq t$,
 - [iii] if $v = 0$, continue to the next j ,
 - [iv] substitute $a = gS$ in W_+ and compute the left kernel of the resulting matrix,
 - [v] if the kernel has dimension 2, or if it has dimension 1 and is generated by a row K with $K_{\rho+1}K_{\rho+2} = 0$, continue to the next σ ,
 - [vi] add the normalized form of $u := -K_{\rho+1}/K_{\rho+2}$ to \mathcal{C} ,
 - (7) for each $u = P/Q$ in $\mathcal{C} \setminus R_\lambda$ do
 - (α) if $\deg P > B_{\text{num}}$ or if $\deg Q > B_{\text{den}}$, then continue to the next σ ,
 - (β) if P/Q cancels the left-hand side of the Riccati equation (R), then add u to R_λ , else continue to the next σ ,
 - (8) continue to step (C)(i) in Algorithm 4, thus quitting the loop over σ .
-

Algorithm 5: Body of step (C)(h) in Algorithm 4.

p. 179) and the definition of a radical provide the first equality in

$$\sqrt{\langle \Sigma \rangle} = \sqrt{\bigcap_{i \in I} \mathfrak{q}_i} = \bigcap_{i \in I} \mathfrak{q}_i,$$

where, additionally, the second is because the ideals \mathfrak{q}_i are prime, therefore radical, and because an intersection of radical ideals is radical ((Zariski and Samuel 1958, Theorem 9, p. 147) or (Cox, Little, and O’Shea 2015, Proposition 16, p. 197)). Finally, any irredundant primary decomposition of $\sqrt{\langle \Sigma \rangle}$ is obtained by retaining a subfamily of the \mathfrak{q}_i , because the latter are prime. This ends the proof as the \mathfrak{q}_i are defined by linear polynomials. \square

Theorem 7.16. *Algorithm 4 terminates and correctly computes all solutions of (R) in $\overline{\mathbb{K}}(x^{1/*})$.*

Proof. The general structure of the algorithm is a loop over λ at step (C), with independent calculations for different λ , so it is sufficient to prove that for each $\lambda \in \Lambda$ the calculation terminates and computes all solutions u with leading coefficient λ .

For each λ , the algorithm introduces the operator $L_\lambda \in \mathbb{K}[\lambda][x]\langle M \rangle$ at step (C)(c) and a basis (z_1, \dots, z_t) of its solutions in $\mathbb{K}[\lambda][[x]]$ at step (C)(d). Note that by Lemma 3.5, the dimension of formal power series solutions of L_λ in $\overline{\mathbb{K}}[[x]]$ is also equal to t , and that the $\mathbb{K}[\lambda]$ -basis z of solutions in $\mathbb{K}[\lambda][[x]]$ is also a $\overline{\mathbb{K}}$ -basis of the solutions in $\overline{\mathbb{K}}[[x]]$. It is therefore sufficient to prove that step (C)(h) computes all rational solutions that can be written $M(az^T)/(az^T)$ for $a \in \mathbb{P}^{t-1}(\overline{\mathbb{K}})$ to get that the algorithm determines all ramified rational solutions $u \in \overline{\mathbb{K}}(x^{1/*})$ of (R) with leading coefficient λ at step (C)(i).

With respect to termination, step (C) may quit early at step (C)(e), but if the run goes beyond step (C)(e), then the basis z contains nonzero entries. By the comment for the case $\rho^{[\infty]} = 2t$ after Lemma 7.4, this forces $\rho^{[\infty]} < 2t$ and the algorithm continues with an unbound inner loop over σ at step (C)(h). This inner loop can quit early at step (C)(h)(4), ending the calculation for the current λ . The inner loop can only be relaunched from step (C)(h)(6)(ε)[i] if a nonlinear polynomial is detected, from step (C)(h)(6)(ε)[v] if a nonzero rational function cannot be defined, and from steps (C)(h)(7)(α) and (β) if a false solution is detected.

At the construction of z at step (C)(d), no z_i is a $O(x^{\sigma_0})$, and therefore no z_i is ever a $O(x^\sigma)$ after the extension step (C)(h)(1). Consequently, ρ is never $2t$ in the inner loop, which is why only the cases $0 \leq \rho \leq 2t - 1$ are considered in the algorithm.

The objects computed inside the loop at step (C)(h), most of which have coefficients in $\mathbb{K}[\lambda]$, relate to the theoretical objects defined in §7.1 and §7.2 for $\mathbb{L} = \overline{\mathbb{K}}$. We already proved that z is a basis for the solutions of L_λ in $\overline{\mathbb{K}}[[x]]$. At step (C)(h)(2), the computation of the minimal basis, by Derksen’s algorithm (1994), Beckermann and Labahn’s algorithm (Beckermann and Labahn 1994) or other known algorithms, is independent of any extension of $\mathbb{K}[\lambda]$ in which the coefficients of the input would be seen. Consequently, the matrix so obtained provides a minimal basis of the approximate syzygy module both over $\mathbb{K}[\lambda][x]$ and over $\overline{\mathbb{K}}[x]$. By the second interpretation, the matrix W computed at step (C)(h)(3) is the matrix W of §7.1 for $\mathbb{L} = \overline{\mathbb{K}}$, and its height, ρ in the algorithm description, is the rank $\rho^{[\sigma]}$ of the theory, also for $\mathbb{L} = \overline{\mathbb{K}}$. For the rest of the proof, we consider the objects $\mathcal{T}^{[\sigma]}$, $\mathcal{A}^{[\sigma]}$, $\mathcal{V}^{[\sigma]}$, $\Sigma^{[\sigma]}$, which are all obtained only from W , and the theory applied to $\mathbb{L} = \overline{\mathbb{K}}$.

Suppose that the algorithm runs without exiting the inner loop, thus making σ grow indefinitely. In particular, ρ is never 0, which would cause early quitting, so we indefinitely have $1 \leq \rho \leq 2t - 1$. From some point on, we have $\mathcal{T}^{[\sigma]} = \mathcal{T}^{[\infty]}$, so that Lemma 7.7 applies, and $\mathcal{V}^{[\sigma]} = \mathcal{A}^{[\sigma]} = \mathcal{A}^{[\infty]}$ holds. If the calculation enters

step (C)(h)(5), it continues to step (C)(h)(7). Otherwise, the calculation enters step (C)(h)(6). Then $\rho \leq 2t-2$ and the ideal $\sqrt{\langle \Sigma^{[\sigma]} \rangle}$ computed at step (C)(h)(6)(γ) is the ideal of the variety $\mathcal{V}^{[\sigma]}$. At this point, it is legitimate to apply Theorem 7.15 to the operator L_λ , which obviously has coefficients in $\overline{\mathbb{K}}$, all other hypotheses being fulfilled in terms of objects over $\mathbb{L} = \overline{\mathbb{K}}$. We conclude that the prime ideals \mathfrak{p}_j have linear generators, so that the algorithm passes beyond step (C)(h)(6)(ε)[i]. By Lemma 7.10, the kernel computed at step (C)(h)(6)(ε)[iv] has dimension 1, and by Lemma 7.7(1), all K computed satisfy $K_{\rho+1}K_{\rho+2} \neq 0$, so that the algorithm also passes beyond step (C)(h)(6)(ε)[v]. In this case again, the calculation continues to step (C)(h)(7). At this point, because $\mathcal{V}^{[\sigma]} = \mathcal{A}^{[\infty]}$ can only describe true solutions, owing to Lemma 7.4, the validation step (C)(h)(7) finds no false solution and the loop ends, a contradiction.

We have just seen that the loop over σ terminates, and that it does after proving that all candidates are true solutions. Consider the final value of σ and exclude the trivial case of early termination at step (C)(h)(4), that is, assume $\rho^{[\sigma]} > 0$. In the inner loop body, the algorithm runs either step (C)(h)(5) or (6). In the first case, $\mathcal{V}^{[\sigma]} = \overline{\mathbb{K}}^t$, which, in view of Lemma 7.8, leads to a rational candidate u parametrized by the free parameter (g_1, \dots, g_t) at step (C)(h)(5)(β). In the second case, the prime decomposition computed at step (C)(h)(6)(γ) represents $\sqrt{\langle \Sigma^{[\sigma]} \rangle}$, so the union of the varieties of the \mathfrak{p}_j in $\overline{\mathbb{K}}^t$ is the cone $\mathcal{V}^{[\sigma]}$, and the rational candidates are the fractions P/Q implied by Lemma 7.11. By our hypothesis of a terminating σ , all the candidate rational solutions obtained previously are verified to be true solutions by step (C)(h)(7), that is $\mathcal{V}^{[\sigma]} \subset \mathcal{A}^{[\infty]}$.

Because of the inclusions $\mathcal{A}^{[\infty]} \subset \mathcal{A}^{[\sigma]} \subset \mathcal{V}^{[\sigma]}$ that are valid independently of σ , we have thus proved $\mathcal{V}^{[\sigma]} = \mathcal{A}^{[\infty]}$, that is, step (C)(h)(8) returns exactly the set of all solutions u with leading coefficient λ . \square

Remark 7.17. Although solving by computing a prime decomposition may exhibit a bad worst-case behavior in theory, we have not investigated how to improve the theoretical complexity of the step after observing that this is never the bottleneck of execution: the computation of Hermite–Padé approximants always takes more time. In addition, the special structure of a union of spaces makes it plausible that one could develop a better method than the general primary decomposition algorithm.

Example 7.18. Let us return to Example 7.14 to show that the number of steps before obtaining the final value of σ can be made arbitrarily large in Algorithm 4. The algorithm first determines the parameters $\nu = \omega/(b-1)$ and $\mu = \omega b/(b-1)$. For $\lambda = \lambda_0$, it then gets $q_{\lambda_0} = b-1$, and the rightmost λ_0 -admissible edge happens to be the leftmost, with slope $-\nu$, thus providing $p_{\lambda_0} = \omega$ and $c_{\lambda_0} = \mu$. This implies $\nu_{\lambda_0} = \mu_{\lambda_0} = 0$. The order $\sigma = \sigma_0$ defined at step (c) for the series expansion at the first iteration of the loop (h) is $\omega + 1$, and solving L_{λ_0} yields $t = 1$ at step (d), with a basis z of the form $z = (z_1)$ for $z_1 = 1 + O(x^{\sigma_0})$. The algorithm verifies that the candidate obtained from $z_1 = 1$ is a false solution. In view of (7.14), to distinguish z_1 from 1, the algorithm will need to increase the order to approximately $(b^r - b)\omega$. For exposition sake, consider a variant setting of the algorithm that multiplies σ by b at each iteration (instead of the golden ratio), so that $\sigma = b^k \sigma_0$ (starting with $k = 0$). The algorithm then needs to reach $k = r$ before z can be distinguished from 1. For $k = r$, if $\sigma - e > B_{\text{num}}$ for the exponent $e = (b^r - b)\omega$ appearing in (7.14), then the rank $\rho^{[\sigma]}$ is zero and the algorithm immediately stops; if $\sigma - e \geq B_{\text{num}}$, the

algorithm continues to increase k until it decides to stop. In all cases, the final value of σ is at least $b^r(\omega + 1)$. For a numerical example, set $b = 4$, $L = x^{10} - M + M^4$. It takes five iterations and computation with accuracy $O(x^{2816})$ to conclude that the Riccati equation has no solution: this is a case where the algorithm stops at $k = r$, with $2816 = 4^4(10 + 1) > B_{\text{num}} = 157$ (for the bound B_{num} relative to L_{λ_0}).

7.5. A supplementary remark on the rank of relations module. A striking fact is that the module of relations contains but is not limited to the direct sum over similarity classes of the module of relations contributed by each class.

The following lemma quantifies the relations that appear in each similarity class.

Lemma 7.19. *For an $\mathbb{L}(x)$ -similarity class $\mathfrak{H}_{\neq 0}$ of $\mathbb{L}(x)$ -hypergeometric elements, let (y_1, \dots, y_s) denote a basis of the \mathbb{L} -vector space \mathfrak{H} and S the column vector $(y_1, \dots, y_s, My_1, \dots, My_s)^T$. The module of the row vectors $R \in \mathbb{L}[x]^{2s}$ such that $RS = 0$ has rank $2s - 1$.*

Proof. Making $F = \mathbb{L}(x)$ in Lemma 2.7 implies that \mathfrak{H} is an \mathbb{L} -vector space. Since the y_i , $1 \leq i \leq s$, lie in the same $\mathbb{L}(x)$ -similarity class, we get equalities $y_i = q_i y_1$ for some q_i in $\mathbb{L}(x)$ and all i . As y_1 is $\mathbb{L}(x)$ -hypergeometric, there exists u_1 in $\mathbb{L}(x)$ such that $My_1 = u_1 y_1$, from which it results $My_i = (Mq_i)u_1 y_1$. Therefore, $y_1, \dots, y_s, My_1, \dots, My_s$ are $2s$ elements of the line $\mathbb{L}(x)y_1$. The module of their linear relations thus has rank $2s - 1$. \square

However, other relations may be produced by the interaction between several similarity classes. We cannot quantify the phenomenon, and merely give an example.

Example 7.20. Let us consider, with $b = 2$, the lcm (least common left multiple) in $\mathbb{Q}[x]\langle M \rangle$ of the operators $L_1 = (1 - 2x^2)M - (1 - 2x)$ and $L_2 = (1 - 3x^2)M - (1 - 3x)$, which respectively annihilate the rational functions $y_1 = 1/(1 - 2x)$ and $y_2 = 1/(1 - 3x)$. This is the operator

$$P_1 := (6x^8 - 5x^4 + 1)M^2 - (6x^6 + 6x^5 + x^4 - x^3 - 4x^2 + x + 1)M + (6x^4 + x^3 - 4x^2 + x).$$

We then slightly modify it by truncating its coefficients, to obtain the operator

$$P_2 := (-5x^4 + 1)M^2 - (x^4 - 5x^3 - 4x^2 + x + 1)M + (x^3 - 4x^2 + x),$$

in such a way that P_2 has a 2-dimensional space of power series solutions as P_1 . One of them is the infinite product

$$y_3 = \prod_{k \geq 0} M^k \frac{1 - 5x^2}{1 - 4x + x^2}$$

and a linearly independent other one is

$$y_4 = x + 5x^2 + 19x^3 + 71x^4 + 265x^5 + 983x^6 + 3667x^7 + 13661x^8 + O(x^9).$$

At this point, we can observe that the lcm L of P_1 and P_2 admits three right factors of order 1, namely $M - u_i$, $1 \leq i \leq 3$, with

$$u_1 = \frac{1 - 2x}{1 - 2x^2}, \quad u_2 = \frac{1 - 3x}{1 - 3x^2}, \quad u_3 = \frac{1 - 4x + x^2}{1 - 5x^2}.$$

As y_1 and y_2 are in the same $\mathbb{L}(x)$ -similarity class, but not y_3 , we could (wrongly) expect a dimension $(2 \cdot 2 - 1) + (2 \cdot 1 - 1) = 4$ for the rank of the syzygy module of the column vector $(y_1, \dots, y_4, My_1, \dots, My_4)^T$. However the operator P_2 factors as

$$P_2 = (1 - 5x^4)(M - v)(M - u_3), \quad \text{with } v = x \frac{1 - 5x^2}{1 - 5x^4}$$

and the operator $M - v$ admits the rational solution $z = x/(1 - 5x^2)$. As a consequence the series y_4 , which satisfies $P_2 y_4 = 0$, also satisfies the equation $(M - u_3)y_4 = cz$ for some constant c . But z , as a rational function, lies in the same similarity class as y_1 and y_2 . So we have an additional relation, between y_1 , y_4 , and My_4 , and the rank of the relations module proves to be 5.

7.6. Rational solving of the linear Mahler equation⁴. In (CDDM 2018), we developed an algorithm to compute the rational solutions of the linear Mahler equation (L). There, we derived degree bounds for the numerators and denominators of rational solutions (see Propositions 3.16, 3.17, and 3.21 in that reference, or, for alternative bounds, (Bell and Coons 2017)). Here we sketch how the approach by Hermite–Padé approximants of the present §7 adapts, based on these bounds, to give an alternative algorithm for computing rational solutions in the linear case.

To this end, we modify the definition of the cone $\mathcal{A}^{[\sigma]}$ in (7.9) and that of the augmented matrix $W_+^{[\sigma]}$ obtained by stacking (7.10). We start with the vector of series $(z_1, \dots, z_t, 1)^T = (z, 1)^T$ and, for any given σ , we compute a minimal basis from which we extract the rows compatible with the degree bounds. This yields a $\rho^{[\sigma]} \times (t + 1)$ -matrix $W^{[\sigma]}$ satisfying

$$W^{[\sigma]}(z, 1)^T = O(x^\sigma).$$

We define $W_+^{[\sigma]}$ as the matrix $W^{[\sigma]}$ augmented at its bottom with the two-row matrix

$$(7.17) \quad \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 & \dots & a_t & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

and we let

$$(7.18) \quad \mathcal{A}^{[\sigma]} = \{0\} \cup \{a \in (\mathbb{L}^t)_{\neq 0} \mid \exists P \in \mathbb{L}[x]_{\neq 0}, \exists Q \in \mathbb{L}[x]_{\neq 0}, \exists \Lambda \in \mathbb{L}[x]^{\rho^{[\sigma]}}, \\ (\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, P, -Q)W_+^{[\sigma]} = 0\}.$$

The motivation is that for $a \in \mathcal{A}^{[\sigma]}$ and corresponding (Λ, P, Q) , if $W^{[\sigma]}(z, 1)^T = 0$, we get $(\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, P, -Q)W_+^{[\sigma]}(z, 1)^T = 0$, implying $Pa z^T - Q = 0$, that is, $Q/P = a z^T$.

In this context, the rank of $W^{[\sigma]}$ satisfies $\rho^{[\sigma]} \leq t$. Again, we distinguish two cases:

- If $\rho^{[\sigma]} = t$, we expect, provided σ is large enough, that (L) admits a full basis of rational solutions. We proceed as in §7.2.1 to prove $\mathcal{A}^{[\sigma]} = \mathbb{L}^t$ and to identify a candidate rational vector solution $K^T \in \mathbb{L}(x)^{t+1}$. If $W^{[\sigma]}(z, 1)^T = 0$, then K must be proportional to $(z, 1)^T$ in $\mathbb{L}(x)^{t+1}$, hence each series z_i is equal to the rational series K_i/K_{t+1} . We therefore consider a parametrized candidate in the form $Q/P = \sum_{i=1}^t g_i(K_i/K_{t+1})$ for the parameter $a = g$ in \mathbb{L}^t .

⁴We are indebted to Alin Bostan for asking us a question that led to the present section.

- Else, $\rho^{[\sigma]} \leq t - 1$, and we proceed as in §7.2.2. The cone $\mathcal{A}^{[\sigma]}$ is given by the vanishing of all minors of rank $\rho^{[\sigma]} + 2$ in $W_+^{[\sigma]}$. (These are the minors of rank $\rho^{[\sigma]} + 1$ in the submatrix obtained by removing the last row and last column.) Extracting coefficients with regard to x yields a linear system in a , so the cone $\mathcal{A}^{[\sigma]}$ is a vector space that can be parametrized in the form $a = gS$. Solving for a left kernel after this specialization delivers a basis of rows of the form $(\Lambda_1, \dots, \Lambda_{\rho^{[\sigma]}}, P, -Q)$, each parametrized by $g \in \mathbb{L}^v$. If $W^{[\sigma]}(z, 1)^T = 0$, then for each row $Paz^T = Q$, so P and Q can be zero only simultaneously: but in this case, the rank of $W^{[\sigma]}$ makes Λ be zero, so the kernel has dimension 1. (In an actual calculation, a failure of the computed kernel to have dimension 1 proves that σ must be increased.) From this, we obtain a parametrized candidate solution Q/P . Cramer’s rules applied to a selection of $\rho^{[\sigma]} + 1$ columns of $W_+^{[\sigma]}$ including the last one show that P is independent of a and Q is linear in a .

We continue as in the Riccati case: all candidates have to be checked against degree bounds and exact evaluation to 0 of the linear Mahler equation, and false solutions require to increase σ .

Experimentally, the algorithm above seems to behave better than our Algorithm 9 in (CDDM 2018). It would thus be of interest to analyze its complexity.

8. IMPLEMENTATION AND BENCHMARK

8.1. Implementation. To test the examples listed in §8.2, we used Dumas’s package `dcfun`⁵. This contains an implementation of Algorithms 1, 3, and 4, and the needed parts of (CDDM 2018) in the computer-algebra system Maple, with a specific call to Singular’s routine for primary decomposition (see below).

The implementation of Algorithms 1 and 3 corresponds to their specification, but the implementation of Algorithm 4 has limitations, which have however no impact on the validity of the treatment of examples in §8.3, as we now explain.

Although algorithms for computing Hermite–Padé approximants allow algebraic coefficients, no implementation was available in Maple beyond rational number coefficients. For step (C)(h)(2) in Algorithm 4, we have therefore used the Maple command `MahlerSystem` in the package `MatrixPolynomialAlgebra` that is restricted to coefficients in \mathbb{Q} , thus forcing $\mathbb{K} = \mathbb{Q}$, and limiting the search of rational solutions of (R) for series in $\bar{\mathbb{Q}}(x^{1/*})$ with leading coefficient λ in \mathbb{Q} . In principle, this makes the implementation able to find solutions corresponding to $\lambda \in \Lambda \cap \mathbb{Q}$, but in practice, all of our examples have $\Lambda \subset \mathbb{Q}$.

Although algorithms for computing primary decompositions allow the search for decompositions over the algebraic closure of \mathbb{K} , implicitly making algebraic extensions as needed along their process, Maple’s command `PrimeDecomposition` in the package `PolynomialIdeals` limits the algebraic numbers used to an algebraic field specified as part of the input. To the best of our knowledge, the only general implementation is available in the computer-algebra system Singular, as the command `absPrimdecGTZ`. The Maple implementation of Algorithm 4 transparently calls Singular at its step (C)(h)(6)(γ).

⁵available from <https://mathexp.eu/dumas/dcfun/>

8.2. Examples. To validate and exemplify our theory, we propose two kinds of examples. The first consists of generating functions of automatic sequences or more generally of sequences satisfying a linear recurrence with constant coefficients of the divide-and-conquer type. The second kind is provided by operators that are lcm of operators vanishing on simple expressions, like rational or power functions. In the subsequent tables and discussions, all examples are referred to by self-explanatory pseudonyms, like `Baum_Sweet` for the first one below.

8.2.1. Generating functions of automatic sequences and like sequences.

Example 8.1 (Baum–Sweet sequence). The Baum–Sweet sequence $(a_n)_{n \in \mathbb{N}}$ is the automatic sequence defined by $a_n = 1$ if the binary representation of n contains no blocks of consecutive 0 of odd length, and $a_n = 0$ otherwise (OEIS, A086747). The generating function $y(x) = \sum_{n \in \mathbb{N}} a_n x^n$ satisfies the Mahler equation

$$M^2 y + x M y - y = 0 \quad (b = 2).$$

Example 8.2 (Rudin–Shapiro sequence). The Rudin–Shapiro sequence $(a_n)_{n \in \mathbb{N}}$ is the automatic sequence defined by $a_n = (-1)^{e_n}$ where e_n is the number of (possibly overlapping) blocks 11 in the binary representation of n (OEIS, A020985). It is characterized by the recurrence relations

$$a_0 = 1, \quad a_{2n} = a_n, \quad a_{2n+1} = (-1)^n a_n,$$

hence its generating function satisfies the Mahler equation

$$2xM^2 y - (x-1)M y - y = 0 \quad (b = 2).$$

Example 8.3 (Stern–Brocot sequence). The Stern–Brocot sequence $(a_n)_{n \in \mathbb{N}}$ (OEIS, A002487) was introduced by Stern to enumerate the nonnegative rational numbers bijectively by the numbers a_n/a_{n+1} . Allouche and Shallit (1992) showed that the sequence is completely defined by $a_0 = 0$, $a_1 = 1$, and for all $n \in \mathbb{N}$,

$$a_{2n} = a_n, \quad a_{4n+1} = a_n + a_{2n+1}, \quad a_{4n+3} = 2a_{2n+1} - a_n.$$

We will re-obtain the well-known fact that its generating function is the Mahler hypergeometric function

$$y(x) = x \prod_{k \geq 0} (1 + x^{2^k} + x^{2^{k+1}}) = \sum_{n \geq 1} a_n x^n \in \mathbb{Z}[[x]],$$

which is obviously a solution of $L_2 y(x) = xy(x) - (1+x+x^2)y(x^2) = 0$ with $b = 2$.

To this end, we follow a method implicit in (Christol, Kamae, Mendès France, and Rauzy 1980) that was detailed in (Allouche 1987). Write $y_1(x) = y(x)$ and introduce $y_2(x) = \sum_{n \in \mathbb{N}} a_{2n+1} x^n$, to obtain

$$\mathbf{y}(x) = A(x)\mathbf{y}(x^2) \quad \text{for} \quad \mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}, \quad A(x) = \begin{pmatrix} 1 & x \\ 1-x & 1+2x \end{pmatrix}.$$

After introducing the Mahler operator for the radix $b = 2$, we get $\mathbf{y} = A(M\mathbf{y}) = A(MA)(M^2\mathbf{y})$, hence, after setting $R = (1, 0)$,

$$y_1 = RA(MA)(M^2\mathbf{y}), \quad M y_1 = R(MA)(M^2\mathbf{y}), \quad M^2 y_1 = R(M^2\mathbf{y}).$$

Because the row vectors $RA(MA)$, $R(MA)$, R are linearly dependent over $\overline{\mathbb{Q}}(x)$, there is an operator L'_2 of order 2 cancelling y_1 . Performing the procedure using

Maple, we readily obtain it in explicit form, so that the generating series $y(x)$ is a solution of

$$L'_2 y(x) = xy(x) - (1 + x + 2x^2)y(x^2) + (1 + x^2 + x^4)y(x^4) = 0.$$

Both Algorithm 3 and Algorithm 4 compute the hypergeometric solutions of L'_2 , to prove that the only ones are multiples of $y(x)$. Correspondingly, one can also verify the relation $L'_2 = (1 - M)L_2$.

Because $y(x)$ is a solution of a linear Mahler equation for radix $b = 2$, for any $k \geq 2$, it is also a solution of a linear Mahler equation for radix b^k that one can make explicit. This gives rise to related operators: an annihilating operator of the generating function $y(x)$, of order 2 for radix $b = 4$, is

$$L'_4 = (1 + x + 2x^2)(1 + x + x^2)^2(1 - x + x^2)^2(1 - x^2 + x^4)^2(1 - x^4 + x^8)M^2 - c_1(x)M + x^3(1 + x^4 + 2x^8) \quad (b = 4)$$

with

$$c_1(x) = \sum_{n=0}^{14} a_{n+1}x^n = 1 + x + 2x^2 + x^3 + 3x^4 + 2x^5 + 3x^6 + x^7 + 4x^8 + 3x^9 + 5x^{10} + 2x^{11} + 5x^{12} + 3x^{13} + 4x^{14}.$$

As well, $y(x)$ is Mahler hypergeometric with respect to radix $b = 4$, as reflected by the factorization $L'_4 = ((2x^2 + x + 1)M - (2x^8 + x^4 + 1))L_4$ for

$$L_4 = (x^2 + x + 1)(x^4 + x^2 + 1)M - x^3 \quad (b = 4).$$

The operators L'_2 and L'_4 give rise to the respective examples `Stern_Brocot_b2` and `Stern_Brocot_b4` in our list.

Example 8.4 (Missing digit in ternary expansion). The sequence 0, 1, 3, 4, ... of nonnegative integers whose ternary expansion does not contain the digit 2 (OEIS, A005836) has a generating function $y(x) = x + 3x^2 + 4x^3 + 9x^4 + 10x^5 + \dots$ annihilated by the operator (used in `no_2s_in_3_exp`)

$$L = x - (1 + 3x + 4x^2)M + 3(1 + x^2)^2M^2 \quad (b = 2).$$

Using either of our algorithms, we find that L admits the unique right-hand factor $M - 1/(3(1 + x))$, corresponding to hypergeometric series solutions of L that are scalar multiples of $(\ln x)^{\log_2(1/3)}/(1 - x)$. This proves that the generating series $y(x)$ is not hypergeometric.

Example 8.5 (Dilcher–Stolarsky formal power series). Inspired by (Dilcher and Stolarsky 2007, Prop. 5.1), we consider the formal power series $F(x) \in \mathbb{Z}[[x]]$ that is a solution of

$$x^4 M^2 y(x) - (1 + x + x^2)M y(x) + y(x) = 0 \quad (b = 4), \quad y(0) = 1.$$

Example 8.6 (Katz–Linden generating function). Katz and van der Linden (2021, §3.1) define a sequence $(A_t(x))_{t \in \mathbb{N}}$ whose generating function $y(w, x) = \sum_{t \geq 0} A_t(x)w^t$ satisfies an order 4 and degree 14 equation $Ly = 0$ with respect to x

with $b = 2$ and

$$\begin{aligned}
L = & -x(x+1)(8x^4w^2 + 4x^2w - x^2 + 2w - 1) \\
& -x(8w^3x^7 - 8w^3x^6 + 8w^3x^5 + 8w^3x^4 - 4w^2x^5 - 4x^4w^2 - wx^5 - 2w^2x^3 - 3wx^4 \\
& \quad + 2w^2x^2 - 4wx^3 + x^4 + 2w^2x - 4x^2w + x^3 + 2w^2 - 3wx + x^2 - w + x)M \\
& \quad + x^2w(16w^3x^8 + 8w^2x^8 + 32w^3x^5 + 4wx^7 + 16w^3x^4 - x^7 + 12x^4w^2 - 4wx^5 \\
& -x^6 + 16w^2x^3 - 4wx^4 + x^5 + 16w^2x^2 + x^4 + 8w^2x - x^3 + 4w^2 - x^2 - x - 1)M^2 \\
& \quad + 2x^4w^2(-8w^2x^{10} - 4wx^9 + 32w^3x^6 - 2wx^8 + x^9 + x^8 - 4wx^4 - 8wx^3 \\
& \quad \quad - 8x^2w + 2x^3 - 4wx + 2x^2 - 2w + x + 1)M^3 \\
& \quad - 8x^{12}w^3(8w^2x^2 + 4wx + 2w - x - 1)M^4.
\end{aligned}$$

The theory we developed in the previous sections made the hypothesis that \mathbb{K} is a computable subfield of \mathbb{C} mostly in order to reuse results from our previous article. However, the theory easily adapts to equations that depend rationally on auxiliary parameters, and our implementation is able to deal with this example as well. It will be called `Katz_Linden` in what follows.

Example 8.7 (Parities and ternary expansion). Adamczewski and Faverjon (2017, Example 8.2) introduce the four generating series $y_i(x) = \sum_{n \geq 0} a_{i,n}x^n$ where $a_{i,n} \in \{0,1\}$ is 1 if and only if the 3-adic expansion of n contains: an even number of 1 and an even number of 2 for $i = 1$; an odd number of 1 and an even number of 2 for $i = 2$; an even number of 1 and an odd number of 2 for $i = 3$; an odd number of 1 and an odd number of 2 for $i = 4$. After showing

$$\mathbf{y}(x) = A(x)\mathbf{y}(x^3) \quad \text{for} \quad \mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ y_3(x) \\ y_4(x) \end{pmatrix}, \quad A(x) = \begin{pmatrix} 1 & x & 0 & x^2 \\ x & 1 & x^2 & 0 \\ 0 & x^2 & 1 & x \\ x^2 & 0 & x & 1 \end{pmatrix},$$

they prove that the $y_i(x)$ are linearly independent over $\overline{\mathbb{Q}}(x)$.

Here, we apply the same method as for Example 8.3, that is, we continue by introducing the Mahler operator for the radix $b = 3$ to get $\mathbf{y} = A(M\mathbf{y}) = A(MA)(M^2\mathbf{y}) = A(MA)(M^2A)(M^3\mathbf{y}) = A(MA)(M^2A)(M^3A)(M^4\mathbf{y})$, hence, after setting $R = (1, 0, 0, 0)$,

$$\begin{aligned}
y_1 &= RA(MA)(M^2A)(M^3A)(M^4\mathbf{y}), & My_1 &= R(MA)(M^2A)(M^3A)(M^4\mathbf{y}), \\
M^2y_1 &= R(M^2A)(M^3A)(M^4\mathbf{y}), & M^3y_1 &= R(M^3A)(M^4\mathbf{y}), & M^4y_1 &= R(M^4\mathbf{y}).
\end{aligned}$$

Because the row vectors $R(M^jA) \cdots (M^3A)$ for $j = 0, \dots, 4$ are linearly dependent over $\overline{\mathbb{Q}}(x)$, there is an operator L of order 4 cancelling y_1 . Performing the procedure using Maple, we readily obtain L with order $r = 4$ and degree $d = 258$, and, repeating the construction (with different R), we observe that the same L actually cancels all y_i . This makes an example that we name `Adamczewski_Faverjon`.

In roughly one second, our Algorithm 4 finds that the only rational solutions to the Riccati Mahler equation are

$$\frac{1}{1-x-x^2}, \quad \frac{1}{1+x-x^2}, \quad \frac{g_1+g_2x^3}{g_1+g_2x} \frac{1}{1+x^2+x^4}.$$

(Algorithm 3 has an equivalent output in over 70 minutes.) This corresponds to hypergeometric solutions of L of the respective forms:

$$\begin{aligned} &\lambda(1 - x - x^2 - x^3 + x^4 + x^5 - x^6 + x^7 + x^8 - x^9 + \dots), \\ &\lambda(1 + x - x^2 + x^3 + x^4 - x^5 - x^6 - x^7 + x^8 + x^9 + \dots), \\ &\lambda(1 + x^2 + x^4 + x^6 + x^8 + \dots) + \mu(x + x^3 + x^5 + x^7 + x^9 + \dots). \end{aligned}$$

Comparing with $y_1(x) = 1 + x^4 + x^8 + \dots$, $y_2(x) = x + x^3 + x^9 + \dots$, $y_3(x) = x^2 + x^6 + \dots$, and $y_4(x) = x^5 + x^7 + \dots$, we deduce that no y_i is not hypergeometric.

8.2.2. Cooked-up examples.

Example 8.8 (Annihilator related to rational functions). We produced two examples as lclm of q first-order operators in $\mathbb{Q}[x]\langle M \rangle$, thus forcing the Riccati equation to have known rational solutions. In both cases, the Riccati equation has an isolated rational solution plus a family parametrized by a projective line. However, by adjusting a coefficient, we forced the number m of distinct λ in the logarithmic parts of solutions of (L) to be either 1 or 2. These examples are named `lclm_⟨q⟩rat_⟨m⟩log`.

For more involved examples, we proceeded as just described to obtain a first operator L_1 , which we then tweaked into an operator L_2 by discarding monomials above some line of a given slope s in the Newton diagram of L_1 . This way, we predict the same dimensions of series solutions, but we expect to lose their hypergeometric nature. After taking the lclm of L_1 and L_2 , we obtain the operators for our examples `lclm_⟨q⟩rat_trunc_sl⟨s⟩`. In particular, the operator in Example 7.20 is the one for example `lclm_2rat_trunc_sl1` in the tables. The corresponding L_1 is of order 2 with all its solutions rational.

Example 8.9 (Annihilator of power functions). For various pairs (b, q) given by a radix $b \leq 5$ and some positive integer $q \leq 5$, we considered some operator $L_{b,q} \in \mathbb{Q}[x]\langle M \rangle$ annihilating the q powers $x^{1/1}, x^{1/2}, \dots, x^{1/q}$. The corresponding example is called `lclm_⟨q⟩pow_b⟨b⟩` in the tables. To obtain the operator $L_{b,q}$, we constructed it as the lclm of binomial annihilators for the relevant $x^{1/i}$, which, to ensure no ramification in these operators, they need not be of order 1. As an example, for $(b, q) = (3, 4)$, we used the annihilators B_i of $x^{1/i}$ given as

$$B_1 = M - x^2, \quad B_2 = M - x, \quad B_3 = M^2 - x^2M, \quad B_4 = M^2 - x^2.$$

Their lclm $L_{3,4}$ has order $r = 6$ and degree $d = 727$.

Example 8.10 (Random equations). To test the robustness of the Hermite–Padé approach, we considered random operators constructed as follows. Given a radix $b \in \{2, 3\}$ and some degree parameter δ , we first draw random A, B , and C , each of the form $VM - U$ for dense polynomials U and V of degree δ in x having integer coefficients in the range $[-1000, 1000]$. Let δ' be the degree in x of $C' := \text{lclm}(A, B)$, normalized to have no denominator. Then $\tilde{C} := C' + x^{\delta'}(M^2 + M + 1)$ has the same dimension of series solutions and series solutions with the same possible valuations as C' , but no more hypergeometric solutions. Then, our random operator is chosen to be $L := \text{lclm}(C, \tilde{C})$. It has a hypergeometric solution, and the Riccati Mahler equation admits the rational function U/V corresponding to C as a rational solution. These are the examples `rmo_b_δ` in the tables, for $b = 2, 3$ and $\delta = 1, \dots, 5$.

8.2.3. Operators related to criteria of differential transcendence. The criteria to be discussed in §9 will lead us to solve the auxiliary Riccati equation (9.2) beside the equations (R), a.k.a. (9.1), for operators of order $r = 2$. We thus introduced problems named `dft_⟨p⟩` where p ranges over relevant problems already listed in §8.2.1.

8.3. Discussion of the timings. We have executed our algorithms on the operators of §8.2 on a Dell Precision Mobile 7550 with i9 processor and 64 GB of RAM, installed with updated Archlinux. For reproducibility of the timings, we have put the processor in a state where thermal status and number of concurrent jobs has no influence,⁶ effectively fixing the cpu frequency to 2.4 GHz, and run all examples one after another. We killed any example above either 12 hours of calculation or 60 GB or used memory.

Table 1 compares our variant Mahler analogues of Petkovšek’s method. It shows the speedup of the improved Algorithm 3, which uses the various prunings and optimizations discussed in §6.5, over the basic Algorithm 1. All those calculations were obtained for the simpler case of the field $\mathbb{L} = \mathbb{Q}$. The reduction of the number of triples (B, A, ζ) considered by the algorithms (typically by a factor in the range 2–10 in our list of examples) induces part of the speedup. In addition, we observe that computing the ancillary operators \tilde{L} and the corresponding sets $Z(\tilde{L})$ takes a significant part of the time in some of the runs of the basic variant (more than half of the total time in cases like `Stern_Brocot_b4` and `dft_Rudin_Shapiro`), and this time is saved by the improved variant. For computations of nonnegligible times, the overall speedup is typically of a few units or dozens of units. We speculate that significant improvements would also occur for calculations with a more general number field \mathbb{L} .

An example will suggest that Algorithms 1 and 3 are slower if we consider $\mathbb{L} = \overline{\mathbb{Q}}$ instead of $\mathbb{L} = \mathbb{Q}$, because the introduction of absolute factorizations of the polynomials ℓ_0 and ℓ_r induces more factors and exponentially more divisors (A, B) to test. Namely, when changing $\mathbb{L} = \mathbb{Q}$ to $\mathbb{L} = \overline{\mathbb{Q}}$, the example `Stern_Brocot_b4` leads to a ramification $q_{\overline{\mathbb{Q}}} = 1$, after which the number of factors of $\ell_0 = (2x^8 + x^4 + 1)x^3$ increases from 4 to 11 (when counted with multiplicities), and that of $\ell_2 = (2x^2 + x + 1)(x^8 - x^4 + 1)(x^2 + x + 1)^2(x^2 - x + 1)^2(x^4 - x^2 + 1)^2$ from 8 to 26, and the number of pairs (A, B) to be potentially tested is changed from $2 \cdot 4 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 3 = 864$ to $2^8 \cdot 4 \cdot 2^2 \cdot 2^8 \cdot 3^2 \cdot 3^2 \cdot 3^4 = 6879707136$: even if the filtering coprimality condition makes the count of pairs effectively leading to a calculation be smaller (that is, the value 57 of ‘tpl’ is smaller than 864), we expect that many more pairs are used over $\overline{\mathbb{Q}}$.

Table 2 compares the improved Mahler analogue of Petkovšek’s method by Algorithm 3 over $\mathbb{L} = \mathbb{Q}$ with the Hermite–Padé approach by Algorithm 4 (which by design is necessarily over $\overline{\mathbb{Q}}$). Even though Algorithm 4 computes a more complete solution set, it is by far the faster algorithm, at least if we exclude the special examples `rmo_b_δ`. Discriminating the solutions over $\mathbb{L} = \mathbb{Q}$ from its complete solution set could be done easily, so solving the Riccati equation over $\mathbb{L} = \mathbb{Q}$ reduces to solving it over $\mathbb{L} = \overline{\mathbb{Q}}$: this makes Algorithm 4 be the better algorithm. For

⁶Using the “performance cpufreq governor” and forbidding any adaptive cpu overclocking, a.k.a. cpu “turbo mode” avoids variations of timings up to a factor of 2.

longer calculations, speedups of Algorithm 4 over Algorithm 3 can be very high (e.g., `Adamczewski_Faverjon`, `dft_Dilcher_Stolarsky`, `dft_Stern_Brocot_b4`).

Examples `rmo_b_delta` were constructed in Example 8.10 to have solutions in $\mathbb{Q}(x)$ and show a situation where Algorithm 4 fails: for example, for `rmo_3_3`, the polynomials ℓ_0 and ℓ_3 have few factors but factors of large degrees; specifically, they have the following factorization patterns:

$$\ell_0 = x^{10}(x^3 + \cdots)(x^{21} + \cdots)(x^{93} + \cdots), \quad \ell_3 = (x^{27} + \cdots)(x^{31} + \cdots)(x^{69} + \cdots).$$

This makes Algorithms 1 and 3 iterate over few pairs (A, B) . In contrast, even if the degree bound (equations (5.1) and (5.2)) and the induced σ needed in Algorithm 4 are not too large, of the order of a few hundreds, the sequences of coefficients of the series solutions are not automatic (see §8.2.1) and involve very large numbers, which dramatically slows down the calculation.

9. DIFFERENTIAL TRANSCENDENCE OF MAHLER FUNCTIONS

Mahler equations originate in number theory, where they were introduced by Mahler as he developed his eponymous method to construct new transcendental numbers; see (Adamczewski 2017) for a recent survey. Consider again the linear Mahler equation (L), this time with $\mathbb{K} = \overline{\mathbb{Q}}$, as well as some solution series $f \in \overline{\mathbb{Q}}[[x]]$. It is classical (Nishioka 1996) that f has a positive radius of convergence and that it can be extended to a meromorphic function on the open unit disk, having a natural boundary on the circle of radius 1, thus being transcendental, unless f is rational.

Concerning values of f , Philippon (2015) proved that for all $\alpha \in \overline{\mathbb{Q}}$ satisfying the relation $|\alpha| < 1$ and such that $\alpha^{b^{\mathbb{N}}}$ does not intersect the zero set of $\ell_0 \ell_r$, the algebraic relations between $f(\alpha), \dots, f(\alpha^{b^{r-1}})$ over $\overline{\mathbb{Q}}$ are specialization at $x = \alpha$ of algebraic relations between the functions $f, \dots, M^{r-1}f$ over $\overline{\mathbb{Q}}(x)$. Therefore, if the latter functions have no algebraic relations, most of their algebraic values are algebraically independent. This could be seen as an analogue of the Hermite–Lindemann theorem, which states that for all nonzero algebraic number α , $\exp(\alpha)$ is transcendental.

The algebraic relations between $f(\alpha), \dots, \partial^n f(\alpha)$, where ∂ denotes the derivation with respect to x , can be studied by the same approach, at least for convenient $\alpha \in \overline{\mathbb{Q}}$ (Adamczewski, Dreyfus, and Hardouin 2021, Theorem 1.5). Their result is that such relations come from specializations of algebraic relations between f and its derivatives, and otherwise that f is differentially algebraic if there exist $n \in \mathbb{N}$ and a nonzero $P \in \overline{\mathbb{Q}}[x][X_0, \dots, X_n]$, such that $P(f, \partial f, \dots, \partial^n f) = 0$. We say that f is differentially transcendental otherwise. This motivates the question of determining whether f is differentially transcendental.

In a way that reinforces the dichotomy between rational and transcendental solutions of (L), Adamczewski, Dreyfus, and Hardouin (2021) recently proved that f is differentially transcendental unless it is a rational function $f \in \overline{\mathbb{Q}}(x)$. This shows that, to prove the differential transcendence of f , it is sufficient to check that it is not rational. To this end, our Algorithm 9 in (CDDM 2018), or for that matter our new development in §7.6, can be used.

In special situations, precisely when $r = 2$ and ℓ_0/ℓ_2 is a monomial, earlier results based on difference Galois theory give to a stronger statement, leading to a criterion for the differentially algebraic independence of $\{f, Mf\}$ and not just of $\{f\}$.

example	br	d	var	tpl	anc	cfs	pol	#	#	tot
Baum_Sweet	22	1	BP	1	0.04	0.00	0.01	0	0	0.06
			IP	1	0.04	0.00	0.02	0	0	0.07
Rudin_Shapiro	22	1	BP	2	0.04	0.00	0.01	0	0	0.07
			IP	1	0.00	0.00	0.02	0	0	0.08
no_2s_in_3_exp	22	4	BP	6	0.06	0.01	0.03	2	1	0.11
			IP	4	0.01	0.00	0.04	2	1	0.12
Stern_Brocot_b2	22	4	BP	4	0.08	0.01	0.02	2	1	0.13
			IP	2	0.01	0.00	0.03	1	1	0.12
Stern_Brocot_b4	42	26	BP	57	19	2.2	1.3	2	1	23
			IP	30	4.7	0.00	0.53	1	1	5.4
Dilcher_Stolarsky	42	4	BP	3	0.06	0.01	0.02	0	0	0.11
			IP	1	0.01	0.00	0.02	0	0	0.09
Katz_Linden	24	14	BP	54	4.2	1.1	12	0	0	17
			IP	8	0.45	0.00	1.5	0	0	2.1
Adamczewski_Faverjon	34	258	BP	168	1307	275	20792	12	3	22412
			IP	92	67	0.00	475	11	3	543
lclm_3rat_1log	33	121	BP	1628	1693	222	7754	21	2	9702
			IP	116	62	0.00	140	5	2	203
lclm_3rat_2log	33	122	BP	1598	1438	184	7762	21	2	9411
			IP	116	67	0.00	147	5	2	215
lclm_2rat_trunc_s10	24	56	BP	1580	3339	391	1432	19	2	5263
			IP	653	270	0.00	219	14	2	490
lclm_2rat_trunc_s11	24	61	BP	2069	7463	1052	2595	9	2	11346
			IP	915	468	0.00	358	8	2	828
dft_Baum_Sweet	42	6	BP	5	0.08	0.02	0.06	0	0	0.18
			IP	1	0.01	0.00	0.03	0	0	0.10
dft_Rudin_Shapiro	42	7	BP	941	83	13	18	0	0	152
			IP	81	3.0	0.00	2.3	0	0	5.8
dft_Stern_Brocot_b2	42	24	BP	50	12	1.0	0.87	2	1	15
			IP	21	2.2	0.00	0.57	1	1	3.0
dft_no_2s_in_3_exp	42	20	BP	98	23	2.6	1.7	2	1	29
			IP	80	8.1	0.00	1.3	2	1	9.6
dft_Dilcher_Stolarsky	162	50	BP							>12 hr
			IP	2025	1064	0.00	2158	0	0	3382
dft_Stern_Brocot_b4	162	348	BP							>12 hr
			IP	1528	26096	0.00	3481	1	1	29670

- ‘var’ stands for the used variant, ‘BP’ (basic) or ‘IP’ (improved).
- ‘tpl’ counts the triples (B, A, ζ) considered (and dealt with) by the loops.
- ‘anc’ is the sum over (B, A) of the times to compute the ancillary operators \tilde{L} .
- ‘cfs’ is, in the basic variant, the sum over (B, A) of the times to compute the sets $Z(\tilde{L})$ of coefficients ζ in solutions u ; it reduces to the single calculation of $Z(L)$ in the improved variant.
- ‘pol’ is the sum over (B, A, ζ) of the times to solve for the polynomials C .
- ‘#’ is the cumulative count over (B, A, ζ) of obtained parametrized solutions u , with possible redundancy.
- ‘#’ counts the number of parametrizations of solutions u retained after removing repeated and embedded parametrizations.
- ‘tot’ is the total time of the ramified rational solving.

(All times measured in seconds.)

Table 1: Comparison of both variants of our Mahler analogue of Petkovšek’s method given by Algorithms 1 and 3 for the search of rational solutions, with $\mathbb{L} = \mathbb{Q}$.

example	br	d	IP		HP					
			tot	fst	dim	σ	rfn	syz	sng	tot
Baum_Sweet	22	1	0.07	0.07	(1, 1)	(6, 6)	0.03	0.03	-	0.13
Rudin_Shapiro	22	1	0.08	0.07	(1, 0)	(6, -)	0.02	0.01	-	0.10
no_2s_in_3_exp	22	4	0.12	0.08	(1, 1)	(33, 9)	0.03	0.08	-	0.21
Stern_Brocot_b2	22	4	0.12	0.07	(1)	(21)	0.01	0.02	-	0.12
Stern_Brocot_b4	42	26	5.4	0.08	(1)	(63)	0.02	0.11	-	0.22
Dilcher_Stolarsky	42	4	0.09	0.07	(2)	(27)	0.04	0.08	0.02	0.23
Katz_Linden	24	14	2.1	0.12	(0, 1, 0, 0)	(-, 69, -, -)	0.06	0.39	-	0.57
Adamczewski_Faverjon	34	258	543	0.16	(4)	(163)	0.32	1.8	0.05	2.4
lclm_3rat_1log	33	121	203	0.08	(3)	(140)	0.16	2.5	0.03	2.9
lclm_3rat_2log	33	122	215	0.09	(2, 1)	(88, 52)	0.07	0.51	-	0.71
lclm_2rat_trunc_sl0	24	56	490	0.11	(4)	(294)	2.6	12	0.05	14
lclm_2rat_trunc_sl1	24	61	828	0.12	(4)	(519)	13	104	0.05	117
lclm_3rat_trunc_sl1	35	1260	>12 hr	0.49	(3, 2)	(574, 268)	11	51	0.07	63
lclm_4pow_b2	27	107	25351	0.20	(1, 4)	(429, 739)	0.16	2.4	-	2.8
lclm_4pow_b3	36	727	>12 hr	0.56	(1, 4)	(108, 174)	0.47	0.64	-	1.7
lclm_4pow_b4	45	989	>12 hr	0.23	(4)	(223)	0.40	0.59	-	1.4
lclm_4pow_b5	55	3103	>12 hr	2.0	(1, 4)	(44, 289)	2.8	0.94	-	5.9
lclm_5pow_b4	47	17270	>60 GB	39	(1, 5)	(274, 1326)	64	6.5	-	115
dft_Baum_Sweet	42	6	0.10	0.08	(2)	(77)	0.06	0.18	0.02	0.37
dft_Rudin_Shapiro	42	7	5.8	0.06	(1, 0)	(88, -)	0.03	0.15	-	0.25
dft_Stern_Brocot_b2	42	24	3.0	0.09	(1)	(59)	0.03	0.10	-	0.22
dft_no_2s_in_3_exp	42	20	9.6	0.09	(1, 1)	(85, 33)	0.07	0.84	-	1.0
dft_Dilcher_Stolarsky	162	50	3382	0.10	(2)	(666)	0.25	3.7	-	4.1
dft_Stern_Brocot_b4	162	348	29670	0.13	(1)	(239)	0.14	2.0	-	2.4
rmo_2_1	23	19	5.3	0.07	(3)	(263)	1.1	23853	0.03	23854
rmo_3_1	33	37	14	0.07	(3)	(133)	0.22	1166	0.03	1167
rmo_2_2	23	44	15							>12 hr
rmo_3_2	33	82	39	0.08	(3)	(247)	2.6	11031	0.03	11034
rmo_2_3	23	69	26							>12 hr
rmo_3_3	33	127	70							>12 hr
rmo_2_4	23	94	41							>12 hr
rmo_3_4	33	172	109							>12 hr
rmo_2_5	23	119	58							>12 hr
rmo_3_5	33	217	166							>12 hr

- ‘tot’ is the total time for ramified rational solving using the improved Mahler analogue of Petkovšek’s approach (IP) or the Hermite–Padé approach (HP).
- ‘fst’ is the time for a first series computation, sufficient to determine the dimensions of series-solutions spaces behind the various logarithmic parts in solutions, provided in the column ‘dim’.
- ‘dim’ is a list, indexed by the $\lambda \in \Lambda$, of the dimension of series appearing in front of $(\ln x)^{\log_b \lambda}$ in solutions.
- ‘ σ ’ is a list with same indexing of the last value of σ used to find the hypergeometric series solutions of L_λ (or ‘-’ when the dimension for λ is 0).
- ‘rfn’ is the cumulative time over λ for all refined series computations up to the corresponding final approximation orders in ‘ σ ’.
- ‘syz’ is the total time for computing minimal bases.
- ‘sng’ is the cumulative time over λ for all prime decompositions computed by calling Singular, or ‘-’ if no prime decomposition was needed for the operator L .

(All times measured in seconds.)

Table 2: Comparison of the improved Mahler analogue of Petkovšek’s method (Algorithm 3, with $\mathbb{L} = \mathbb{Q}$) and the Hermite–Padé approach (Algorithm 4, over \mathbb{Q}).

Dreyfus, Hardouin, and Roques (2018) proved a criterion for differential algebraicity: if the difference Galois group of (\mathbf{L}) contains $\mathrm{SL}_r(\overline{\mathbb{Q}})$, then the nonzero series solutions are differentially transcendental⁷. Furthermore, when ℓ_0/ℓ_r is a monomial, $f, \dots, M^{r-1}f$ are differentially algebraically independent, that is, for all $n \in \mathbb{N}$, there is no algebraic relation over $\overline{\mathbb{Q}}(x)$ between the $\partial^j M^i(f)$, $0 \leq j \leq n$, $0 \leq i \leq r-1$. Note that we cannot expect a better result with respect to r since by (\mathbf{L}) , $f, \dots, M^r(f)$ are linearly dependent over $\overline{\mathbb{Q}}(x)$. When $r = 2$, the following theorem provides a criterion in terms of the nonexistence of solutions of Riccati equations for the Galois group to contain $\mathrm{SL}_2(\overline{\mathbb{Q}})$.

Theorem 9.1 (Roques (2018, §6)). *For $r = 2$, assume the existence of a nonzero solution $f \in \overline{\mathbb{Q}}[[x]]$ of (\mathbf{L}) . Assume further that the following equations have no solutions in $\overline{\mathbb{Q}}(x^{1/*})$:*

$$(9.1) \quad \ell_2 u M u + \ell_1 u + \ell_0 = 0,$$

$$(9.2) \quad u M^2 u + \left(M^2 \left(\frac{\ell_0}{\ell_1} \right) - M \left(\frac{\ell_1}{\ell_2} \right) + \frac{\ell_2}{\ell_1} M \left(\frac{\ell_0}{\ell_2} \right) \right) u + \frac{\ell_2 \ell_0 M \ell_0}{\ell_1^2 M \ell_1} = 0.$$

Then, the difference Galois group of (\mathbf{L}) contains $\mathrm{SL}_2(\overline{\mathbb{Q}})$.

It is worth mentioning that (9.2) can be viewed as a Riccati equation in radix b^2 . Additionally, the Galoisian criterion in (Dreyfus, Hardouin, and Roques 2018) and the previous theorem straightforwardly combine into the following corollary.

Corollary 9.2. *For $r = 2$, assume the existence of a nonzero solution $f \in \overline{\mathbb{Q}}[[x]]$ of (\mathbf{L}) . Assume further that (9.1) and (9.2) have no solutions in $\overline{\mathbb{Q}}(x^{1/*})$. Then, f is differentially transcendental. If we further assume that ℓ_0/ℓ_2 is a monomial, we obtain that f and Mf are differentially algebraically independent.*

Note that the first part of this corollary is a criterion for differential transcendence now obsoleted by (Adamczewski, Dreyfus, and Hardouin 2021): in the latter work, differential transcendence is obtained without any assumption either on the differential Galois group or on Riccati equations, just as a consequence of the linear Mahler equation having no ramified rational solution. On the other hand, the second part of the corollary is a stronger but conditional result, which requires such assumptions.

We tested our implementation on our 6 examples of order $r = 2$ appearing in §8.2. Calculations by Algorithm 4 took a total of 8.4 seconds: a total of 1.0 seconds for (9.1); a total of 8.3 seconds for (9.2). In the worse case, solving of (9.2) takes 18 times as much time as the solving of (9.1). We also executed Algorithm 3, which could also do all 6 examples, in much more time (9.4 hours in total). With either algorithm, this has the status of a proof. The stronger criterion can also be applied to 3 of the 6 examples: `Baum_Sweet`, `Rudin_Shapiro`, and `Dilcher_Stolarsky`. Our calculations proved the differentially algebraic independence of $\{f, Mf\}$ in these three cases. (In the 3 remaining cases, no independence can be concluded.)

⁷In Dreyfus, Hardouin, and Roques (2018), the series have coefficients in \mathbb{C} but everything remains correct if we replace \mathbb{C} by the algebraically closed field $\overline{\mathbb{Q}}$.

REFERENCES

- Abramov, Sergei A., Peter Paule, and Marko Petkovšek (Feb. 1998). “ q -Hypergeometric solutions of q -difference equations”. In: *Discrete Mathematics*. Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics 180.1, pp. 3–22. DOI: [10.1016/S0012-365X\(97\)00106-4](https://doi.org/10.1016/S0012-365X(97)00106-4). URL: <https://www.sciencedirect.com/science/article/pii/S0012365X97001064> (visited on 01/04/2024).
- Abramov, Sergei A. and Marko Petkovšek (1995). “Finding all q -hypergeometric solutions of q -difference equations”. In: *Proc. FPSAC '95*.
- Adamczewski, Boris (2017). “Mahler’s method”. In: *Doc. Math.* Extra Vol. Mahler Selecta, pp. 95–122. DOI: [10.25537/dm.2019.SB-95-122](https://doi.org/10.25537/dm.2019.SB-95-122).
- Adamczewski, Boris, Thomas Dreyfus, and Charlotte Hardouin (2021). “Hypertranscendence and linear difference equations”. In: *J. Amer. Math. Soc.* 34.2, pp. 475–503. DOI: [10.1090/jams/960](https://doi.org/10.1090/jams/960). URL: <https://doi.org/10.1090/jams/960>.
- Adamczewski, Boris and Colin Faverjon (2017). “Méthode de Mahler : relations linéaires, transcendance et applications aux nombres automatiques”. In: *Proc. Lond. Math. Soc. (3)* 115.1, pp. 55–90. URL: <https://doi.org/10.1112/plms.12038>.
- Allouche, Jean-Paul (1987). “Automates finis en théorie des nombres”. In: *Exposition. Math.* 5.3, pp. 239–266.
- Allouche, Jean-Paul and Jeffrey Shallit (1992). “The ring of k -regular sequences”. In: *Theoret. Comput. Sci.* 98.2, pp. 163–197.
- Becker, Thomas and Volker Weispfenning (1993). *Gröbner Bases*. Springer New York. DOI: [10.1007/978-1-4612-0913-3](https://doi.org/10.1007/978-1-4612-0913-3).
- Beckermann, Bernhard and George Labahn (1994). “A uniform approach for the fast computation of matrix-type Padé approximants”. In: *SIAM J. Matrix Anal. Appl.* 15.3, pp. 804–823. DOI: [10.1137/s0895479892230031](https://doi.org/10.1137/s0895479892230031).
- (2000). “Fraction-free computation of matrix rational interpolants and matrix GCDs”. In: *SIAM J. Matrix Anal. Appl.* 22.1, pp. 114–144. DOI: [10.1137/s0895479897326912](https://doi.org/10.1137/s0895479897326912).
- Beke, Emanuel (1894). “Die Irreducibilität der homogenen linearen Differentialgleichungen”. In: *Mathematische Annalen* 45.2, pp. 278–294.
- Bell, Jason P. and Michael Coons (2017). “Transcendence tests for Mahler functions”. In: *Proc. Amer. Math. Soc.* 145.3, pp. 1061–1070.
- Bendixson, Ivar (1892). “Sur les équations différentielles linéaires homogènes”. In: *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar* 49, pp. 91–105.
- Bostan, Alin, Tanguy Rivoal, and Bruno Salvy (July 2023). *Minimization of differential equations and algebraic values of E -functions*. DOI: [10.48550/arXiv.2209.01827](https://doi.org/10.48550/arXiv.2209.01827). arXiv: [2209.01827v3](https://arxiv.org/abs/2209.01827v3) [cs, math]. URL: <http://arxiv.org/abs/2209.01827v3> (visited on 07/25/2023).
- Bronstein, Manuel and Thom Mulders (n.d. [1999?]). *Rational heuristics for rational solutions of Riccati equations*. URL: <https://citeseerx.ist.psu.edu/pdf/3ef0af0b941a70cb2b68cd6bf2ec93b3c0859cc8>.
- Bronstein, Manuel and Marko Petkovšek (1993). *On Ore rings, linear operators and factorisation*. Tech. rep. ETH Zurich, p. 19. DOI: [10.3929/ETHZ-A-000915149](https://doi.org/10.3929/ETHZ-A-000915149). URL: <http://hdl.handle.net/20.500.11850/68860>.

- Christol, Gilles, Teturo Kamae, Michel Mendès France, and Gérard Rauzy (1980). “Suites algébriques, automates et substitutions”. In: *Bull. Soc. Math. France* 108.4, pp. 401–419.
- Chyzak, Frédéric, Thomas Dreyfus, Philippe Dumas, and Marc Mezzarobba (2018). “Computing solutions of linear Mahler equations”. In: *Math. Comp.* 87.314, pp. 2977–3021.
- Cluzeau, Thomas and Mark Van Hoeij (2004). “A modular algorithm for computing the exponential solutions of a linear differential operator”. In: *Journal of Symbolic Computation* 38.3, pp. 1043–1076.
- Cohn, Richard M. (1965). *Difference algebra*. Vol. 17. Interscience tracts in pure and applied mathematics. John Wiley and Sons.
- Cox, David A., John Little, and Donal O’Shea (May 13, 2015). *Ideals, Varieties, and Algorithms*. Springer-Verlag GmbH. URL: https://www.ebook.de/de/product/23693438/david_a_cox_john_little_donal_o_shea_ideals_varieties_and_algorithms.html.
- Decker, Wolfram, Gert-Martin Greuel, and Gerhard Pfister (1999). “Primary Decomposition: Algorithms and Comparisons”. In: *Algorithmic Algebra and Number Theory*. Ed. by B. Heinrich Matzat, Gert-Martin Greuel, and Gerhard Hiss. Springer, pp. 187–220. DOI: [10.1007/978-3-642-59932-3_10](https://doi.org/10.1007/978-3-642-59932-3_10).
- Derksen, Harm (1994). *An algorithm to compute generalized Padé–Hermite forms*. Report 9403. Dept. of Math., Catholic University Nijmegen.
- Dilcher, Karl and Kenneth B Stolarsky (2007). “A polynomial analogue to the Stern sequence”. In: *Internat. J. Number Theory* 3.01, pp. 85–103.
- Dreyfus, Thomas, Charlotte Hardouin, and Julien Roques (2018). “Hypertranscendence of solutions of Mahler equations”. In: *J. Eur. Math. Soc.* 20, pp. 2209–2238.
- Dumas, Philippe (1993). “Récurrences mahlériennes, suites automatiques, études asymptotiques”. PhD thesis. Université Bordeaux I.
- Fabry, Eugène (1885). “Sur les intégrales des équations différentielles linéaires à coefficients rationnels”. PhD thesis. Paris: Faculté des sciences de Paris. URL: <http://patrimoine.sorbonne-universite.fr/idurl/1/3199>.
- Faverjon, Colin and Marina Poulet (Nov. 2022). “An algorithm to recognize regular singular Mahler systems”. In: *Mathematics of Computation* 91.338, pp. 2905–2928. DOI: [10.1090/mcom/3758](https://doi.org/10.1090/mcom/3758). URL: <https://www.ams.org/mcom/2022-91-338/S0025-5718-2022-03758-3/> (visited on 03/15/2023).
- Faverjon, Colin and Julien Roques (2022). *Hahn Series and Mahler Equations: Algorithmic Aspects*.
- Gianni, Patrizia, Barry Trager, and Gail Zacharias (1988). “Gröbner bases and primary decomposition of polynomial ideals”. In: *J. Symb. Comput.* 6.2, pp. 149–167. DOI: [https://doi.org/10.1016/S0747-7171\(88\)80040-3](https://doi.org/10.1016/S0747-7171(88)80040-3).
- Gosper, R. William (Jan. 1978). “Decision procedure for indefinite hypergeometric summation”. In: *Proceedings of the National Academy of Sciences* 75.1, pp. 40–42. DOI: [10.1073/pnas.75.1.40](https://doi.org/10.1073/pnas.75.1.40). URL: <https://pnas.org/doi/full/10.1073/pnas.75.1.40> (visited on 01/04/2024).
- Hendriks, Peter A. (1997). “An algorithm for computing a standard form for second-order linear q -difference equations”. In: *J. Pure Appl. Algebra* 117–118, pp. 331–352. DOI: [10.1016/s0022-4049\(97\)00017-0](https://doi.org/10.1016/s0022-4049(97)00017-0).

- Hendriks, Peter A. (1998). “An algorithm determining the difference Galois group of second order linear difference equations”. In: *J. Symb. Comput.* 26.4, pp. 445–461. DOI: [10.1006/jsco.1998.0223](https://doi.org/10.1006/jsco.1998.0223).
- Ibarra, Oscar H., Shlomo Moran, and Roger Hui (1982). “A generalization of the fast LUP matrix decomposition algorithm and applications”. In: *J. Algorithms* 3.1, pp. 45–56. DOI: [10.1016/0196-6774\(82\)90007-4](https://doi.org/10.1016/0196-6774(82)90007-4).
- Katz, Daniel J. and Courtney M. van der Linden (2021). *Peak sidelobe level and peak crosscorrelation of Golay–Rudin–Shapiro Sequences*. arXiv: [2108.07318](https://arxiv.org/abs/2108.07318).
- Mahler, K. (1968). “Perfect systems”. In: *Compositio Math.* 19.2, pp. 95–166.
- Markoff, André (1891). “Sur la théorie des équations différentielles linéaires”. In: *Comptes Rendus Hebdomadaires des Séances de l’Académie des Sciences, Paris* 113, pp. 790–791.
- Neiger, Vincent (2016). “Bases of relations in one or several variables: fast algorithms and applications”. PhD thesis. École normale supérieure de Lyon, France. URL: <https://tel.archives-ouvertes.fr/tel-01431413>.
- Nishioka, Kumiko (1996). *Mahler functions and transcendence*. Vol. 1631. Lecture Notes in Mathematics. Springer Verlag, Berlin, pp. viii+185.
- NIST (2010). *Digital Library of Mathematical Functions*. Ed. by F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain. Release 1.1.4. URL: <http://dlmf.nist.gov/>.
- OEIS Foundation Inc. (2021). *The On-Line Encyclopedia of Integer Sequences*. URL: <http://oeis.org>.
- Petkovšek, Marko (1992). “Hypergeometric solutions of linear recurrences with polynomial coefficients”. In: *J. Symb. Comput.* 14, pp. 243–264.
- Pflügel, Eckhard (1997). “An algorithm for computing exponential solutions of first order linear differential systems”. In: *Proceedings of the 1997 international symposium on Symbolic and algebraic computation - ISSAC ’97*. Kihei, Maui, Hawaii, United States: ACM Press, pp. 164–171. DOI: [10.1145/258726.258773](https://doi.org/10.1145/258726.258773). URL: <http://portal.acm.org/citation.cfm?doid=258726.258773> (visited on 07/25/2023).
- Philippon, Patrice (2015). “Groupes de Galois et nombres automatiques”. In: *J. London Math. Soc.* 92.3, pp. 596–614.
- van der Put, Marius and Michael F. Singer (1997). *Galois theory of difference equations*. Springer.
- (2003). *Galois theory of linear differential equations*. Vol. 328. Grundlehren der mathematischen Wissenschaften. Springer. URL: <http://www4.ncsu.edu/~singer/papers/dbook2.ps>.
- Roques, Julien (2018). “On the algebraic relations between Mahler functions”. In: *Trans. Amer. Math. Soc.* 370.1, pp. 321–355.
- (2023). “Frobenius method for Mahler equations”. In: *Journal of the Mathematical Society of Japan*, pp. 1–40. DOI: [10.2969/jmsj/89258925](https://doi.org/10.2969/jmsj/89258925). URL: <https://doi.org/10.2969/jmsj/89258925>.
- Singer, Michael F. (Aug. 2016). “Algebraic and algorithmic aspects of linear difference equations”. In: *Galois theories of linear difference equations: an introduction* (Santa Marta, Columbia). Ed. by Charlotte Hardouin, Jacques Sauloy, and Michael F. Singer. Vol. 211. Mathematical Surveys and Monographs. Proceedings

- of a CIMPA research school. American Mathematical Society, pp. 1–41. DOI: <https://doi.org/10.1090/surv/211>.
- van Hoeij, Mark (1999). “Finite singularities and hypergeometric solutions of linear recurrence equations”. In: *Journal of Pure and Applied Algebra* 139.1, pp. 109–131. DOI: [10.1016/S0022-4049\(99\)00008-0](https://doi.org/10.1016/S0022-4049(99)00008-0).
- Wibmer, Michael (2013). *Algebraic Difference Equations*. Lecture notes. URL: https://drive.google.com/open?id=1KmEURssrRQm_L6EE35XY568XcQqwVLuR.
- Zariski, Oscar and Pierre Samuel (1958). *Commutative algebra, Volume I*. The University Series in Higher Mathematics. With the cooperation of I. S. Cohen. D. Van Nostrand Co., Inc., Princeton, New Jersey, pp. xi+329.

FRÉDÉRIC CHYZAK, INRIA (FRANCE)
Email address: `frederic.chyzak@inria.fr`

THOMAS DREYFUS, INSTITUT DE MATHÉMATIQUES DE BOURGOGNE, UMR 5584 CNRS, UNIVERSITÉ DE BOURGOGNE, F-21000, DIJON, FRANCE
Email address: `thomas.dreyfus@cnrs.fr`

PHILIPPE DUMAS, INRIA (FRANCE)
Email address: `philippe.dumas@inria.fr`

MARC MEZZAROBBA, LIX, CNRS, ÉCOLE POLYTECHNIQUE, INSTITUT POLYTECHNIQUE DE PARIS, 91120 PALAISEAU, FRANCE
Email address: `marc@mezzarobba.net`