ELSEVIER

Theoretical Computer Science 157 (1996) 3–33

Theoretical Computer Science

An introduction to pseudo-linear algebra

Manuel Bronstein[a,*] Marko Petkovšek[b]

[a] *Institute for Scientific Computation, ETH Zentrum, CH-8092 Zürich, Switzerland*
[b] *Department of Mathematics, University of Ljubljana, 61111 Ljubljana, Slovenia*

Abstract

Pseudo-linear algebra is the study of common properties of linear differential and difference operators. We introduce in this paper its basic objects (pseudo-derivations, skew polynomials, and pseudo-linear operators) and describe several recent algorithms on them, which, when applied in the differential and difference cases, yield algorithms for uncoupling and solving systems of linear differential and difference equations in closed form.

0. Introduction

Linear ordinary differential equations are equations of the form

$$a_n(t)\frac{\mathrm{d}^n y(t)}{\mathrm{d}t^n} + \cdots + a_1(t)\frac{\mathrm{d}y(t)}{\mathrm{d}t} + a_0(t)y(t) = b(t) \tag{1}$$

or systems of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} = A(t)\begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} + B(t) \tag{2}$$

while linear ordinary difference equations are equations of the form

$$a_n(t)y(t+n) + \cdots + a_1(t)y(t+1) + a_0(t)y(t) = b(t) \tag{3}$$

or systems of the form

$$\begin{bmatrix} y_1(t+1) \\ \vdots \\ y_n(t+1) \end{bmatrix} = A(t)\begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} + B(t) \tag{4}$$

where in both cases the unknown y and the coefficients are functions of the (continuous or discrete) variable t. Those two types of equations are closely connected, the

[*] Corresponding author. E-mail: bronstein@inf.ethz.ch.

SSDI 0304-3975(95)00173-5

various algorithms for solving or otherwise manipulating them have some interesting similarities [1], and on occasion, methods devised for one type of equation can be used for the other type [15]. A comparison of the algebraic properties of those equations points to the existence of some common mathematical abstraction behind them. This abstraction is provided by *pseudo-linear algebra*, an area of mathematics with origins in the 1930s, whose objects of study are skew polynomials [13], Ore'33 which represent single equations (1,3), and pseudo-linear operators [10] which represent systems (2),(4).

Algebraic algorithms originally developed for differential equations and systems [3, 5, 16–18] have been recently generalized to difference equations [14, 15] and arbitrary pseudo-linear equations [6, 22]. This enables us to present in this paper an algorithmic introduction to pseudo-linear algebra. After introducing the basic objects of study in Section 1, we describe their basic arithmetic operations, followed by a factorisation algorithm for skew-polynomials and a weak Frobenius form for pseudo-linear operators. Those last two algorithms are an important part of computer algebra solvers for equations of the form (1)–(4).

All fields in this paper are commutative, rings are noncommutative unless explicitly stated otherwise, and all rings and fields have characteristic 0.

1. The basic objects

1.1. Pseudo-derivations

Let k be a field and $\sigma : k \longrightarrow k$ be an injective endomorphism of k.

Definition 1. A *pseudo-derivation* w.r.t. σ is any map $\delta : k \longrightarrow k$ satisfying

$$\delta(a+b) = \delta a + \delta b \quad \text{and} \quad \delta(ab) = \sigma(a)\,\delta b + \delta a\, b \quad \text{for any } a, b \in k\,. \tag{5}$$

Example. If $\sigma = 1_k$ then (5) is just the rule for a derivation on k, so the derivations on k are exactly all the pseudo-derivations w.r.t. the identity. The pair (k, δ) is called a differential field in that case.

Example. For any injective endomorphism σ and any $\alpha \in k$, the map $\delta_\alpha = \alpha(\sigma - 1_k)$ given by $\delta_\alpha a = \alpha(\sigma(a) - a)$ is a pseudo-derivation w.r.t. σ. Indeed,

$$\delta_\alpha(a+b) = \alpha(\sigma(a+b) - (a+b)) = \alpha(\sigma(a) - a) + \alpha(\sigma(b) - b) = \delta_\alpha(a) + \delta_\alpha(b)$$

and

$$\delta_\alpha(ab) = \alpha(\sigma(ab) - ab) = \sigma(a)\alpha(\sigma(b) - b) + \alpha(\sigma(a) - a)b = \sigma(a)\,\delta_\alpha b + \delta_\alpha a\, b\,.$$

A pseudo-derivation of that form is called an *inner derivation.*

Example. For any injective endomorphism σ, the zero map δ_0 is an inner derivation, hence a pseudo-derivation w.r.t. k. The pair (k, σ) is called a difference field in that case, the associated difference operator being $\Delta = \delta_1 = \sigma - 1_k$.

The above three examples exhaust all the possible pseudo-derivations over a (commutative) field:

Lemma 1. *Let k be a field, σ an injective endomorphism of k, and δ a pseudo-derivation of k. Then,*

(i) *If $\sigma \neq 1_k$ then there is an element $\alpha \in k$ such that $\delta = \alpha(\sigma - 1_k) = \delta_\alpha$.*

(ii) *If $\delta \neq 0$ then there is an element $\beta \in k$ such that $\sigma = \beta\,\delta + 1_k$.*

Proof. Since k is commutative, $\delta(ab) = \delta(ba)$ for any $a, b \in k$, so applying (5) to both sides gives

$$\sigma(a)\,\delta b + \delta a\, b = \sigma(b)\,\delta a + \delta b\, a$$

and, after rearranging,

$$(\sigma(a) - a)\ \delta b = (\sigma(b) - b)\ \delta a\,. \tag{6}$$

(i) If $\sigma \neq 1_k$ then there is an element $a \in k$ such that $\sigma(a) \neq a$. Let $\alpha = \delta a/(\sigma(a) - a)$. Then it follows from (6) that $\delta b = \alpha\,(\sigma(b) - b)$ for all $b \in k$, hence $\delta = \alpha(\sigma - 1_k)$.

(ii) If $\delta \neq 0$ then there is an element $a \in k$ such that $\delta a \neq 0$. Let $\beta = (\sigma(a) - a)/\delta a$. Then it follows from (6) that $\sigma(b) = \beta\,\delta b + b$ for all $b \in k$, hence $\sigma = \beta\,\delta + 1_k$. □

Definition 2. The *constant subfield* of k (with respect to σ and δ) is

$$\mathrm{Const}_{\sigma,\delta}(k) = \{a \in k \text{ such that } \sigma(a) = a \text{ and } \delta a = 0\}.$$

It is easily checked that $\mathrm{Const}_{\sigma,\delta}(k)$ is a subfield of k, since it is the intersection of two subfields.

1.2. Univariate skew-polynomials

Let k, σ and δ be as in the previous section.

Definition 3 (*Ore* [13]). The left skew polynomial ring given by σ and δ is the ring $(k[x], +, \cdot)$ of polynomials in x over k with the usual polynomial addition, and multiplication given by

$$\underline{xa = \sigma(a)x + \delta a} \quad \text{for any } a \in k. \tag{7}$$

To avoid confusing it with the usual commutative polynomial ring $k[x]$, the left skew polynomial ring is denoted $k[x; \sigma, \delta]$, and its elements are called *skew polynomials* or

Ore polynomials since they were introduced by Ore [13]. The multiplication defined by (7) can be uniquely extended to an associative multiplication on monomials by

$$(ax^n)(bx^m) = (ax^{n-1})(xb)x^m = (ax^{n-1})(\sigma(b)x^{m+1} + \delta b\, x^m) \quad \text{for } n > 0 \tag{8}$$

and to arbitrary polynomials by distributivity:

$$\left(\sum_i a_i x^i\right)\left(\sum_j b_j x^j\right) = \sum_i \sum_j (a_i x^i)(b_j x^j).$$

Let $A, B \in k[x;\sigma,\delta] \setminus \{0\}$, ax^n and bx^m be the leading monomials of A and B respectively, where $a \neq 0$, $b \neq 0$, and $n, m \geqslant 0$. Then, by (8), the leading monomial of AB is $a\sigma^n(b)x^{n+m}$. Since σ is injective, $a\sigma^n(b) \neq 0$, so

$$\deg(AB) = \deg(A) + \deg(B) \geqslant \max(\deg(A), \deg(B)).$$

This equality implies that $k[x;\sigma,\delta]$ has no zero divisors, and that the degree function satisfies the inequality of a Euclidean norm. In fact, $k[x;\sigma,\delta]$ possesses a right Euclidean division algorithm, and a left Euclidean division if σ is an automorphism. Those will be presented in Section 3.

Example. For any differential field k with derivation δ, $k[D;1_k,\delta]$ is the usual ring of linear ordinary differential operators under composition.

Example. If $k = \mathbb{C}(n)$ and σ is the automorphism of k over $\mathbb{C}$ that takes n to $n+1$, then $k[E;\sigma,0]$ is the ring of linear ordinary recurrence operators, while $k[E;\sigma,\Delta]$ is the ring of linear ordinary difference operators where $\Delta = \sigma - 1_k$.

Example. If $k = \mathbb{C}(q)(t)$ and σ is the automorphism of k over $\mathbb{C}(q)$ that takes t to qt, then $k[B;\sigma,\Delta]$ is the ring of linear ordinary q-difference operators where $\Delta = (\sigma - 1_k)/(t(q-1))$.

Example. For any field k, $k[x;1_k,0] \simeq k[x]$ is the commutative ring of the usual polynomials over k. Thus, polynomials are a special case of skew-polynomials.

1.3. Pseudo-linear maps

Let k, σ, δ be as above and V be a vector space over k.

Definition 4 (Jacobson [10]). A map $\theta : V \to V$ is called *k-pseudo-linear* (w.r.t. σ and δ) if $\theta(au) = \sigma(a)\,\theta(u)$ for difference ring

$$\theta(u+v) = \theta u + \theta v \quad \text{and} \quad \theta(au) = \sigma(a)\,\theta u + \delta a\, u \quad \text{for any } a \in k, u, v \in V. \tag{9}$$

Lemma 2. *Any k-pseudo-linear map is* $\mathrm{Const}_{\sigma,\delta}(k)$*-linear.*

Proof. Let $\theta : V \to V$ be k-pseudo-linear, $c \in \mathrm{Const}_{\sigma,\delta}(k)$ and $u, v \in V$. Then,

$$\theta(cu+v) = \theta(cu) + \theta v = (\sigma(c)\,\theta u + \delta c\, u) + \theta v = c\,\theta u + \theta v. \quad \square$$

Suppose now that $\dim_k(V) = n$ is finite, and let $\mathscr{B} = (b_1, \ldots, b_n)$ be a given basis for V over k. Then the *matrix of* θ *w.r.t.* $\mathscr{B}$ is the matrix $M_{\mathscr{B}}(\theta) = (m_{ij})$ with entries in k given by $\theta b_i = \sum_{j=1}^{n} m_{ji} b_j$ for all i's. The action of θ on the coordinates with respect to $\mathscr{B}$ is then given by

$$\theta \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = M_{\mathscr{B}}(\theta)\, \sigma \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \delta \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}. \tag{10}$$

Conversely, for any $n \times n$ matrix M with entries in k, the map defined on V by (10) with $M_{\mathscr{B}}(\theta)$ replaced by M is k-pseudo linear, and its matrix w.r.t $\mathscr{B}$ is M.

For any $P \in GL_n(k)$, $\mathscr{E} = P\mathscr{B}$ is also a basis for V over k, and the matrix of θ w.r.t. $\mathscr{E}$ is given by the following change of basis formula [10]: Jacobson

$$M_{\mathscr{E}}(\theta) = P^{-1} M_{\mathscr{B}}(\theta)\underline{\sigma(P)} + P^{-1}\delta(P), \tag{11}$$

where σ and δ are applied pointwise to P.

Example. Let k be a differential field k with derivation δ, n be a positive integer, and A an $n \times n$ matrix with coefficients in k. Then the map $\theta : k^n \rightarrow k^n$ given by

$$\theta \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \delta y_1 \\ \vdots \\ \delta y_n \end{bmatrix} + A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

is pseudo-linear w.r.t $1_k, \delta$.

Example. Let $k = \mathbb{C}(m)$, σ be the automorphism of k over $\mathbb{C}$ that takes m to $m+1$, and A an $n \times n$ matrix with coefficients in k. Then the map $\theta : k^n \rightarrow k^n$ given by

$$\theta \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sigma y_1 \\ \vdots \\ \sigma y_n \end{bmatrix} + A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

is pseudo-linear w.r.t $\sigma, 0$.

Example. For any field k and any vector space V over k, any k-linear map from V to V is pseudo-linear w.r.t. $1_k, 0$, so linear maps are a special case of pseudo-linear maps. The change of basis formula (11) becomes

$$M_{\mathscr{E}}(\theta) = P^{-1} M_{\mathscr{B}}(\theta) P$$

which is the usual formula for linear maps.

Example. Let $k = \mathbb{C}(x)$, $\sigma = 1_k$, $\delta = \mathrm{d}/\mathrm{d}x$, $\mathscr{B}$ be the canonical basis for k^2 over k, $\theta : k^2 \rightarrow k^2$ be given in the basis $\mathscr{B}$ by

$$\theta \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \delta y_1 \\ \delta y_2 \end{bmatrix} - M \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where

$$M = \frac{1}{x^2}\begin{bmatrix} x & -x^2-1 \\ x^4+2x^2 & -2x^3-x \end{bmatrix}.$$

θ is pseudo-linear with $M_{\mathscr{B}}(\theta) = -M$. Let $\mathscr{E} = P\mathscr{B}$, where

$$P = \frac{1}{x}\begin{bmatrix} 1 & x \\ 2x & x^2 \end{bmatrix}.$$

By (11) the matrix of θ w.r.t. $\mathscr{E}$ is

$$M_{\mathscr{E}}(\theta) = P^{-1}(-M)P + P^{-1}\delta(P) = \begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}.$$

Since the coordinates in terms of $\mathscr{E}$ are given by $z = P^{-1}y$, we get

$$\theta\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \delta z_1 \\ \delta z_2 \end{bmatrix} + M_{\mathscr{E}}(\theta)\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \delta z_1 + xz_1 \\ \delta z_2 + z_1 + xz_2 \end{bmatrix},$$

where

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = P^{-1}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -xy_1 + y_2 \\ 2y_1 - y_2/x \end{bmatrix}.$$

Note that the system of equations $\theta z = 0$ is uncoupled in the basis $\mathscr{E}$.

The previous discussion described how to make skew polynomials act on a vector space given a pseudo-linear transformation. We show now that such a transformation can always be given on k or any field extension of k to which σ and δ can be extended.

Definition 5. We say that a field extension K of k is compatible with k if σ can be extended to an injective field endomorphism of K, and δ can be extended to a pseudo-derivation of K with respect to σ.

Note that k itself is compatible with k, and that if K is a compatible field extension of k, then $\mathrm{Const}_{\sigma,\delta}(k) \subseteq \mathrm{Const}_{\sigma,\delta}(K)$. We can describe all the k-pseudo linear maps on a compatible extension:

Lemma 3. *Let K be a compatible field extension of k. Then, for any $c \in K$, the map $\theta_c : K \to K$ given by*

$$\theta_c a = c\,\sigma(a) + \delta a \tag{12}$$

is K-pseudo-linear. Conversely, for any K-pseudo-linear map $\theta : K \to K$ there is an element $c \in K$ such that $\theta = \theta_c$ as given in (12).

Proof. Let $a, b \in K$. Using (5) and that σ is a field homomorphism, we get

$$\theta_c(a+b) = c\,\sigma(a+b) + \delta(a+b) = c\,(\sigma(a) + \sigma(b)) + \delta a + \delta b = \theta_c a + \theta_c b$$

and

$$\theta_c(ab) = c\,\sigma(ab) + \delta(ab) = c\,\sigma(a)\,\sigma(b) + (\sigma(a)\,\delta b + \delta a\, b) = \sigma(a)\,\theta_c b + \delta a\, b$$

which proves that θ_c is K-pseudo-linear, hence $\mathrm{Const}_{\sigma,\delta}(K)$-linear by Lemma 2.

To prove the converse, note that by the pseudo-linearity of θ,

$$\theta a = \theta(a\,1) = \sigma(a)\,\theta 1 + \delta a\,,$$

hence $\theta = \theta_c$, where $c = \theta 1$. □

2. Skew polynomials as linear operators

We describe in this section how arbitrary skew polynomials can act on vector spaces over k, and thus be viewed as $\mathrm{Const}_{\sigma,\delta}(k)$-linear operators. Throughout this section, let k, σ, δ and $k[x;\sigma,\delta]$ be as in the previous section.

2.1. The linear action

Given a vector space V over k, any k-pseudo-linear map $\theta : V \to V$ induces an action $*_\theta : k[x;\sigma,\delta] \times V \to V$ given by by interpreting x as θ.

$$\left(\sum_{i=0}^{n} a_i x^i\right) *_\theta u = \sum_{i=0}^{n} a_i \theta^i u$$

for any $u \in V$. This action is linear with respect to the constants of k, so the elements of $k[x;\sigma,\delta]$ can be viewed as linear operators acting on V. When there is no ambiguity from the context, we write $*$ instead of $*_\theta$. It turns out that the multiplication in $k[x;\sigma,\delta]$ corresponds to the composition of operators.

Theorem 1. *$(AB) * u = A * (B * u)$ for any $A, B \in k[x;\sigma,\delta]$ and $u \in V$.*

Proof. We first prove by induction on n that

$$(ax^n bx^m) * u = ax^n * (bx^m * u) \tag{13}$$

for any $n, m \geqslant 0$, $a, b \in k$, and $u \in V$. If $n = 0$, then

$$(ax^0 bx^m) * u = (abx^m) * u = ab\,\theta^m u = a * (b\,\theta^m u) = ax^0 * (bx^m * u).$$

Suppose now that $n > 0$ and that (13) holds for $n-1$. Then, using (8) and (9),

$$\begin{aligned}(ax^n bx^m) * u &= \big((ax^{n-1})\,(\sigma(b)\,x^{m+1} + \delta b\, x^m)\big) * u\\ &= (ax^{n-1}\sigma(b)\,x^{m+1}) * u + (ax^{n-1}\delta b\, x^m) * u\\ &= ax^{n-1} * (\sigma(b)\,x^{m+1} * u) + ax^{n-1} * (\delta b\, x^m * u)\\ &= ax^{n-1} * \big(\sigma(b)\,\theta^{m+1}u + \delta b\,\theta^m u\big)\\ &= ax^{n-1} * \theta(b\,\theta^m u) = ax^{n-1} * (x * (bx^m * u))\\ &= ax^n * (bx^m * u).\end{aligned}$$

Writing now $A = \sum_i a_i x^i$ and $B = \sum_j b_j x^j$, we have

$$\begin{aligned}(AB) * u &= \left(\sum_i \sum_j (a_i x^i)(b_j x^j)\right) * u = \sum_i \sum_j ((a_i x^i b_j x^j) * u) \\ &= A * (B * u). \quad \square\end{aligned}$$

2.2. *Zeros of skew-polynomials*

Definition 6. Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo-linear, and $p \in k[x;\sigma,\delta]$. We say that $\alpha \in V$ is a zero of p (w.r.t. θ) if $p *_\theta \alpha = 0$. We also say that p is an annihilator for α (w.r.t. θ) if $p \neq 0$ and $p *_\theta \alpha = 0$.

Since the action of p on V is linear with respect to $C = \mathrm{Const}_{\sigma,\delta}(k)$, it follows that the zeros of p in V always form a vector space over C, hence that $p *_\theta 0 = 0$ for any $p \in k[x;\sigma,\delta]$. It is also clear that if $\alpha \in V$ has an annihilator, then it has (at least) one of minimal degree. Such a skew-polynomial is called a minimal annihilator for α over k.

Lemma 4. *Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo linear, and $p, q \in k[x;\sigma,\delta]$. Then, any zero of q in V is also a zero of pq.*

Proof. By Theorem 1, we have

$$(pq) *_\theta \alpha = p *_\theta (q *_\theta \alpha) = p *_\theta 0 = 0. \quad \square$$

Of course, the converse does not generally hold since the multiplication is not commutative. As for polynomials, however, minimal annihilators divide all the annihilators exactly on the right:

Theorem 2. *Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo linear, $\alpha \in V$ be a zero of some element of $k[x;\sigma,\delta]\setminus\{0\}$, and $q \in k[x;\sigma,\delta]\setminus\{0\}$ be a minimal annihilator for $\alpha \in k$. Then, for any $p \in k[x;\sigma,\delta]$, $p *_\theta \alpha = 0$ if and only if q divides p exactly on the right.*

Proof. Let $q \in k[x;\sigma,\delta]\setminus\{0\}$ be a minimal annihilator for α, and $p \in k[x;\sigma,\delta]$ be such that $p *_\theta \alpha = 0$. Let $p = aq + r$ be the right Euclidean division of p by q, where either $r = 0$ or $\deg(r) < \deg(q)$. Using Theorem 1 we obtain

$$0 = p *_\theta \alpha = (aq) *_\theta \alpha + r *_\theta \alpha = a *_\theta (q *_\theta \alpha) + r *_\theta \alpha = a *_\theta 0 + r *_\theta \alpha = r *_\theta \alpha\,.$$

Since q has minimal degree among the annihilators of α over k, it follows that $r = 0$ and hence that q divides p exactly on the right.

Conversely, if q divides p exactly on the right, then α is a zero of p by Lemma 4. $\square$

As a consequence, any common zero of two or more skew-polynomials must be a zero of their right gcd.

Corollary 1. *Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo linear, $\alpha \in V$ and $p,q \in k[x;\sigma,\delta]$. Then $p *_\theta \alpha = q *_\theta \alpha = 0$ if and only if* $\operatorname{gcrd}(p,q) *_\theta \alpha = 0$.

Proof. Let $g = \operatorname{gcrd}(p,q)$, and suppose first that $p *_\theta \alpha = q *_\theta \alpha = 0$. If $p = q = 0$, then $g = 0$ so $g *_\theta \alpha = 0$. Otherwise, let $r \in k[x;\sigma,\delta]\backslash\{0\}$ be a minimal annihilator for α. Then, r divides p and q exactly on the right, so it divides g exactly on the right, which implies that $g *_\theta \alpha = 0$ by Theorem 2.

Conversely, suppose that $g *_\theta \alpha = 0$. Then, since $p = ag$ and $q = bg$ for some $a,b \in k[x;\sigma,\delta]$, we have $p *_\theta \alpha = q *_\theta \alpha = 0$ by Lemma 4. □

2.3. Hyperexponential zeros

Hyperexponential elements are the pseudo-linear generalisations of hypergeometric sequences or functions with rational logarithmic derivatives. They play a key role in algorithms for solving or factoring pseudo-linear equations.

Definition 7. Let V be a vector space over k, and $\theta : V \longrightarrow V$ be k-pseudo linear. We say that $\alpha \in V$ is hyperexponential over k (w.r.t. θ) if $\alpha \neq 0$ and $\theta\alpha = u\alpha$ for some $u \in k^*$.

Lemma 5. *Let V be a vector space over k, and $\theta : V \longrightarrow V$ be k-pseudo linear. Then, $\alpha \in V$ is hyperexponential over k if and only if $\alpha \neq 0$ and α has an annihilator of the form $ax + b$ for some $a,b \in k^*$.*

Proof. Suppose that $\alpha \in V$ is hyperexponential over k. Then, $\alpha \neq 0$ and $\theta\alpha = u\alpha$ for some $u \in k^*$, which implies that $x - u \in k[x;\sigma,\delta]$ is an annihilator for α.

Conversely, suppose that $\alpha \neq 0$ and that $(ax+b) *_\theta \alpha = 0$ for some $a,b \in k^*$. Then, $\theta\alpha = (-b/a)\alpha \neq 0$, so α is hyperexponential over k. □

Abramov [2] has noted that the hyperexponential elements are in some sense "eigenvectors" of all skew-polynomials:

Theorem 3 (Abramov [2]). *Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo linear, and $\alpha \in V$. If α is hyperexponential over k, then for every $p \in k[x;\sigma,\delta]$, there exists $u_p \in k$ such that $p *_\theta \alpha = u_p\alpha$.*

Proof. Suppose that $\alpha \in V$ is hyperexponential over k. Then, $\alpha \neq 0$ and $\theta\alpha = u\alpha$ for some $u \in k^*$. We first show by induction that for each $n \geqslant 0$, there exists $u_n \in k$ such that $\theta^n\alpha = u_n\alpha$. For $n = 0$, we have $\theta^0\alpha = \alpha$ so $u_0 = 1$. Suppose that $\theta^n\alpha = u_n\alpha$ for some $n \geqslant 0$ and $u_n \in k$. Then,

$$\theta^{n+1}\alpha = \theta\theta^n\alpha = \theta(u_n\alpha) = \sigma(u_n)\theta\alpha + \delta u_n\ \alpha = (\sigma(u_n)u + \delta u_n)\alpha$$

which proves the claim. Writing $p = \sum_{i=0}^{n} a_i x^i$ we get

$$p *_\theta \alpha = \sum_{i=0}^{n} a_i \theta^i \alpha = \left(\sum_{i=0}^{n} a_i u_i \right) \alpha. \qquad \Box$$

As a consequence, hyperexponential solutions of nonhomogeneous equations with hyperexponential right-hand sides must have a very special form:

Corollary 2. *Let V be a vector space over k, $\theta : V \longrightarrow V$ be k-pseudo linear, and $\alpha, \beta \in V$ be hyperexponential over k. If $p *_\theta \beta = \alpha$ for some $p \in k[x; \sigma, \delta]$, then $\beta = u\alpha$ for some $u \in k^*$.*

Proof. Suppose that $p *_\theta \beta = \alpha$ for some $p \in k[x; \sigma, \delta]$. Since β is hyperexponential over k, $p *_\theta \beta = c\beta$ for some $c \in k$ by Theorem 3. Thus $c\beta = \alpha$ which implies that $c \neq 0$, hence that $\alpha = u\beta$ where $u = c^{-1}$. □

We say that $p \in k[x; \sigma, \delta]$ has an hyperexponential solution if there exists a vector space V over k, a k-pseudo-linear map $\theta : V \longrightarrow V$ and $\alpha \in V$ hyperexponential over k such that $p *_\theta \alpha = 0$. The existence of hyperexponential solutions is closely connected to first-order right factors.

Theorem 4. *If $p \in k[x; \sigma, \delta] \setminus \{0\}$ has an hyperexponential solution, then there exists $u \in k^*$ such that $x - u$ divides p exactly on the right.*

Proof. Let α be a hyperexponential solution of p. By Lemma 5, α has an annihilator of the form $ax + b$ for some $a, b \in k^*$. Since $\alpha \neq 0$, $x + b/a$ is a minimal annihilator for α, so it divides p exactly on the right by Theorem 2. □

Thus, if we can find a hyperexponential solution of p, then it must have a right factor of degree 1, something which is independent of the choice of the pseudo-linear operator chosen for the action of $k[x; \sigma, \delta]$. A converse to Theorem 4 would allow one to prove that a skew polynomial does not have a right factor of degree 1. We need an extra hypothesis on our rings for that, namely that we can construct solutions of skew polynomials of degree 1.

Definition 8. Let $c \in k$. We say that $k[x; \sigma, \delta]$ is *c-solvable* if for any $a, b \in k^*$, there exists a compatible extension K of k, and $\alpha \in K^*$ such that $(ax + b) *_{\theta_c} \alpha = 0$, where θ_c is given by (12).

Theorem 5. *Let $c \in k$. If $k[x; \sigma, \delta]$ is c-solvable and $p \in k[x; \sigma, \delta]$ has a right factor of degree* 1, *then either x divides p exactly on the right, or p has an hyperexponential solution w.r.t. θ_c.*

Proof. Suppose that $k[x; \sigma, \delta]$ is c-solvable and that $p = q(ax + b)$ for some $a, b \in k$ with $a \neq 0$. If $b = 0$, then x is a right factor of p. Otherwise, $b \neq 0$, so there exists a compatible extension K of k, and $\alpha \in K^*$ such that $(ax+b) *_{\theta_c} \alpha = 0$. By Lemma 5, this implies that α is hyperexponential over k, and we have

$$p *_{\theta_c} \alpha = (q(ax+b)) *_{\theta_c} \alpha = q *_{\theta_c} (ax+b) *_{\theta_c} \alpha = q *_{\theta_c} 0 = 0\,,$$

so p has a hyperexponential solution w.r.t. θ_c. □

Thus, if x is not a right factor of p and we can produce $c \in k$ such that $k[x;\sigma,\delta]$ is c-solvable and p has no hyperexponential solution w.r.t. θ_c, then p has no right factor of degree 1 in $k[x;\sigma,\delta]$. This fact will be important in the factorisation algorithm.

3. Basic arithmetic

Let $k[x;\sigma,\delta]$ be a skew-polynomial ring, $A, B \in k[x;\sigma,\delta]\setminus\{0\}$, ax^n and bx^m be their leading terms, and suppose that $n \geqslant m$. The right Euclidean division of A by B is performed as follows: let

$$Q_0 = \frac{a}{\sigma^{n-m}(b)}x^{n-m}. \tag{14}$$

The leading monomial of Q_0B is ax^n, so we can recursively divide $A - Q_0B$ by B on the right, obtaining $Q_1, R \in k[x;\sigma,\delta]$ such that $A - Q_0B = Q_1B + R$ and $\deg(R) < m$. We then have

$$A = QB + R\,,$$

where $Q = Q_0 + Q_1$ and $\deg(R) < \deg(B)$. R is called the *right-remainder* of A by B and is denoted $\operatorname{rrem}(A,B)$, while Q is called the *right-quotient* of A by B and is denoted by $\operatorname{rquo}(A,B)$.

If σ is an automorphism of k, then there is a similar left Euclidean division, where we let

$$Q_0 = \sigma^{-m}\left(\frac{a}{b}\right)x^{n-m}$$

and, dividing recursively $A - BQ_0$ by B on the left, obtain $Q, R \in k[x;\sigma,\delta]$ such that

$$A = BQ + R$$

and $\deg(R) < m$. Q and R are called the left-quotient and left-remainder of A by B in that case.

We can also compute the *right* (resp. *left*) *Euclidean remainder sequence* given by $R_0 = A, R_1 = B$ and $R_i = \operatorname{rrem}(R_{i-2}, R_{i-1})$ (resp. $\operatorname{lrem}(R_{i-2}, R_{i-1})$) for $i \geqslant 2$, and the *greatest common right* (resp. *left*) *divisor* of A and B which is the last nonzero element of that sequence.

Example. Let $k = \mathbb{C}(n)$, σ is the automorphism of k over $\mathbb{C}$ that takes n to $n+1$, and $\delta = 0$. We compute the right gcd in $k[E;\sigma,0]$ of

$$A = n(n+1)E^2 - 2n(n-1)E + n^3 - 3n + 2$$

and

$$B = E^2 - (2n+1)E + n^2 - 5\,.$$

We have $R_0 = A$, $R_1 = B$, and by (14), $Q_0 = n(n+1)/\sigma^0(1)x^0 = n(n+1)$ and

$$R_0 - Q_0R_1 = R_0 - n(n+1)R_1 = n(2n^2+n+3)E - n^4 - n^3 + 6n^2 + 2n + 2 = R_2\,.$$

Dividing further on the right we get

$$R_1 = \left(\frac{1}{(n+1)(2n^2+5n+6)}E - \frac{3n^4 + 11n^3 + 26n^2 + 30n + 14}{n(n+1)(2n^2+n+3)(2n^2+5n+6)}\right)R_2 + R_3\,,$$

where

$$R_3 = \mathrm{gcrd}(A,B) = \frac{n^8 + 2n^7 - 4n^6 - 20n^5 + 4n^4 + 38n^3 + n^2 - 2n + 28}{n(n+1)(2n^2+n+3)(2n^2+5n+6)}\,. \tag{15}$$

Since $R_3 \in k^*$, we get by Corollary 1 that the system of difference equations

$$\begin{cases} n(n+1)y(n+2) - 2n(n-1)y(n+1) + (n^3 - 3n + 2)y(n) = 0 \\ y(n+2) - (2n+1)y(n+1) + (n^2 - 5)y(n) = 0 \end{cases}$$

does not have any nonzero solution.

The extended right (resp. left) Euclidean algorithm yield nontrivial common left (resp. right) multiples of A and B:

```
$R_0 \leftarrow A,\ R_1 \leftarrow B$
$A_0 \leftarrow 1,\ A_1 \leftarrow 0$
$B_0 \leftarrow 0,\ B_1 \leftarrow 1 \leftarrow 1$
while $R_i \neq 0$ do
    $i \leftarrow i + 1$
    $Q_{i-1} \leftarrow \mathrm{rquo}(R_{i-2}, R_{i-1})$
    $R_i \leftarrow \mathrm{rrem}(R_{i-2}, R_{i-1})$
    $A_i \leftarrow A_{i-2} - Q_{i-1}A_{i-1}$
    $B_i \leftarrow B_{i-2} - Q_{i-1}B_{i-1}$
$n \leftarrow i.$
```

It is easy to see by induction on i that

$$R_i = A_iA + B_iB \tag{16}$$

and, running induction backwards, that R_{n-1} right-divides R_i for $n \geqslant i \geqslant 0$. It follows that

$$R_{n-1} = A_{n-1}A + B_{n-1}B = \mathrm{gcrd}(A,B)\,.$$

Since $R_n = 0$ (the terminating condition of the *while* loop), we have

$$A_nA + B_nB = 0$$

hence $A_nA = -B_nB$ is a common left multiple of A and B. From the above algorithm we have

$$\deg(R_i) < \deg(R_{i-1}) \quad \text{and} \quad \deg(Q_{i-1}) = \deg(R_{i-2}) - \deg(R_{i-1}) \quad \text{for } 2 \leqslant i \leqslant n.$$

Hence by induction on i, we see that

$$\deg(A_i) = \deg(B) - \deg(R_{i-1}) \quad \text{and} \quad \deg(B_i) = \deg(A) - \deg(R_{i-1})$$

for $2 \leqslant i \leqslant n$. It follows that $\deg(A_n) = \deg(B) - \deg(R_{n-1})$ and $\deg(B_n) = \deg(A) - \deg(R_{n-1})$, so $A_n \neq 0$ and $B_n \neq 0$. Hence $A_nA = -B_nB$ is a nonzero common left multiple of A and B. In fact, it is a *least* such multiple. To see this, assume that $CA = -DB$ is some common left multiple of A and B. Now let (cf. [7, 11, Ex. 4.6.1.18])

```
$C_0 = -D$
$C_1 = C$
  for $i = 2, 3, \ldots, n$ do
    $C_i = C_{i-2} - C_{i-1}Q_{i-1}$.
```

An easy induction on i shows that

$$C_{i-1}R_i - C_iR_{i-1} = 0 \quad C_{i-1}A_i - C_iA_{i-1} = (-1)^iC \quad C_{i-1}B_i - C_iB_{i-1} = (-1)^iD$$

for $1 \leqslant i \leqslant n$. It follows that $C_nR_{n-1} = C_{n-1}R_n = 0$, hence that $C_n = 0$. Therefore, A_n right-divides C, and B_n right-divides D. Thus,

$$A_nA = -B_nB = \operatorname{lclm}(A, B)$$

is a nonzero left least common multiple of A and B. The fact that two nonzero elements have a nonzero left (resp. right) least common multiple is called the *left* (*resp. right*) *Ore condition*, and rings without zero divisors which satisfy this condition are called *left* (*resp. right*) *Ore rings*. We thus have a constructive proof of:

Theorem 6. *$k[x; \sigma, \delta]$ is a left Ore ring. If σ is an automorphism of k, then $k[x; \sigma, \delta]$ is also a right Ore ring.*

Note that skew-polynomial rings can be defined over a ring k rather than a field [8], in which case there are examples of left skew polynomial rings which are not right Ore rings, and that $k[x; \sigma, \delta]$ is a right Ore ring if and only if σ is an automorphism [8, Ex. 0.8.2].

4. Factorisation of skew-polynomials

We begin with the usual definition of an irreducible element, namely one that cannot be broken into a product of two non-units.

Definition 9. $p \in k[x;\sigma,\delta] \setminus k$ is *irreducible* if $p = ab$ for $a, b \in k[x;\sigma,\delta]$ implies that $a \in k$ or $b \in k$.

$r \in k[x;\sigma,\delta]$ is *similar* to $s \in k[x;\sigma,\delta]$ if $\operatorname{lclm}(s,t) = rt$ for some $t \in k[x;\sigma,\delta]$ such that $\operatorname{gcrd}(s,t) \in k^*$.

Note that two similar elements have the same degree, and that any element of $k[x;\sigma,\delta]$ has a factorisation into irreducibles [13]. Such a factorisation is not unique however, as the following example of differential operators from [12] shows: for any $a \in \mathbb{C}$,

$$D^2 - \frac{2}{t}D + \frac{2}{t^2} = \left(D - \frac{1}{t(1+at)}\right)\left(D - \frac{1+2at}{t(1+at)}\right)$$

in $\mathbb{C}(t)[D; 1, \mathrm{d}/\mathrm{d}t]$. However, the following result of Ore shows that any two factorisations into irreducibles are closely related, in particular they have the same number of factors and the same multisets of degrees.

Theorem 7 (Ore [13, Theorem II/1]). *Let $p \in k[x;\sigma,\delta]$ be monic. If $r_1 \cdots r_m$ and $s_1 \cdots s_n$ are two factorisations of p into irreducibles then $m = n$ and the factors are similar in pairs.*

4.1. The generalized Wronskian

Definition 10. Let V be an algebra over k, $\theta : V \to V$ be pseudo-linear, $y_1, \ldots, y_m \in V$, and $n \geqslant m$ be an integer. The nth generalized Wronskian of $(y_1, \ldots, y_m)$ be the $n \times m$ matrix

$$\mathscr{M} = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \\ \theta y_1 & \theta y_2 & \cdots & \theta y_m \\ \theta^{n-1} y_1 & \theta^{n-1} y_2 & \cdots & \theta^{n-1} y_m \end{pmatrix}. \qquad (17)$$

In addition, for any set $S = \{s_1, \ldots, s_m\}$ of m integers with $1 \leqslant s_1 < \ldots < s_m \leqslant n$, we let $\mathscr{M}_S$ be the submatrix obtained from the rows $s_1, \ldots, s_m$ of $\mathscr{M}$, and $[S] = [s_1, \ldots, s_m]$ be its determinant.

Note that $[1, \ldots, m]$ is the Wronskian of $y_1, \ldots, y_m$ in the differential case, and their Casoratian in the difference case. In both of those cases, it is nonzero if and only if $y_1, \ldots, y_m$ are linearly independent over $\mathrm{Const}_{\sigma,\delta}(k)$. It turns out that quotients of any nonzero m by m minors of $\mathcal{M}$ are all in k, which implies that they are hyperexponential in the differential and difference cases.

Lemma 6. *Let V be an algebra over k, $\theta : V \to V$ be k-pseudo-linear, $p \in k[x;\sigma,\delta]$ be of degree $m > 0$, $y_1,\ldots,y_m \in V$ be zeros of p (not necessarily linearly independent over* $\mathrm{Const}_{\sigma,\delta}(k)$*), and $n \geqslant m$ be an integer. Then for any set $S = \{s_1,\ldots,s_m\}$ of m integers with $1 \leqslant s_1 < \ldots < s_m \leqslant n$, there exists $u_S \in k$ such that $[S] = u_S\,[1,\ldots,m]$. Furthermore, if either $\sigma = 1_k$ or $\delta = 0$, then either $[S] = 0$ or $[S]$ is hyperexponential over k.*

Proof. By induction on s_m. If $s_m = m$, then $S = \{1,\ldots,m\}$, so $[S] = [1,\ldots,m]$. Let now N be such that $m \leqslant N < n$ and suppose that the lemma holds for any sorted set S of m integers with $s_m \leqslant N$. Let S be such that $s_m = N+1$, and write $S = S^* \cup \{N+1\}$, where $S^* = \{s_1,\ldots,s_{m-1}\}$ and $s_{m-1} \leqslant N$. Let $x^N = qp + r$ be the right division of x^N by p where $r = \sum_{j=0}^{m-1} u_j x^j$. Then, for each i,

$$\theta^N y_i = x^N *_\theta y_i = q *_\theta (p *_\theta y_i) + r *_\theta y_i = q *_\theta 0 + r *_\theta y_i = \sum_{j=0}^{m-1} u_j \theta^j y_i\,.$$

Since the u_j's do not depend on i, we have $[S] = \det(\mathcal{M}_S) = \sum_{j=0}^{m-1} u_j \det(A_j)$ where the first $m-1$ rows of each A_j are the first $m-1$ rows of $\mathcal{M}_S$ and the mth row of A_j is $(\theta^j y_1,\ldots,\theta^j y_m)$ i.e. the $(j+1)$th row of $\mathcal{M}$. Hence, A_j is a row permutation of $\mathcal{M}_{S_j}$ where $S_j = S^* \cup \{j+1\}$. Since $j+1 \leqslant N$, $\max(S_j) \leqslant N$, so $[S_j] = u_{S_j}[1,\ldots,m]$ for some $u_{S_j} \in k$ by the induction hypothesis. Therefore,

$$[S] = \sum_{j=0}^{m-1} \pm u_j [S_j] = \sum_{j=0}^{m-1} \pm u_j u_{S_j} [1,\ldots,m]$$

which proves the first part of the lemma.

Suppose now that $\sigma = 1_k$ or $\delta = 0$, and that $[S] \neq 0$ for some S. Then $[1,\ldots,m] \neq 0$ which implies that the y_i's form a fundamental set of solutions of $p *_\theta y = 0$. Hence, $\theta([1,\ldots,m])/[1,\ldots,m] \in k$ by either Liouville's relation [6, 17] or its difference analogue [6], so we get

$$\frac{\theta[S]}{[S]} = \frac{\theta u_S}{u_S} + \frac{\theta[1,\ldots,m]}{[1,\ldots,m]} \in k\,. \quad \square$$

Those minors can also be connected to the coefficients of an annihilator, in a role similar to the symmetric functions for the usual polynomials:

Theorem 8. *Let V be an algebra over k, $\theta : V \to V$ be k-pseudo-linear,*

$$p = x^m - \sum_{i=0}^{m-1} a_i x^i \quad \in k[x;\sigma,\delta]$$

where $m > 0$, $y_1, \ldots, y_m \in V$ be zeros of p linearly independent over $\mathrm{Const}_{\sigma,\delta}(k)$, *and $n \geqslant m$ be an integer. If either $\sigma = 1_k$ or $\delta = 0$, then*

$$a_i = (-1)^{m-i+1} \frac{[\overline{i+1}]}{[1,\ldots,m]} \tag{18}$$

for $0 \leqslant i < m$, where $\bar{j} = \{1,\ldots,m+1\} \setminus \{j\}$.

Proof. This is formula (9) of [6] in the differential case, and formula (15) of [6] in the difference case. □

4.2. Outline of the factorisation algorithm

We describe now an algorithm that reduces the problem of factoring in $k[x;\sigma,\delta]$ to the problem of finding all the irreducible right factors of degree 1. We first reduce the problem to one of the differential or difference field case.

Theorem 9. *If $\sigma \neq 1_k$, then $k[x;\sigma,\delta]$ is isomorphic as a left skew polynomial ring to $k[y;\sigma,0]$ via the isomorphism φ_α given by*

$$\varphi_\alpha\left(\sum_i a_i x^i\right) = \sum_i a_i \left(\frac{y + \delta\alpha}{\alpha - \sigma(\alpha)}\right)^i,$$

where α is any element of k such that $\alpha \neq \sigma(\alpha)$.

Proof. This is a special case of Proposition 3.1 of [8, §8.3]. □

As a consequence, either we are in the differential operator case where $\sigma = 1_k$, or factoring in $k[x;\sigma,\delta]$ is equivalent to factoring in $k[y;\sigma,0]$, which is the difference case. Hence, we can assume in the rest of this section that either $\sigma = 1_k$ or that $\delta = 0$. Since $k[x;\sigma,\delta]$ is 0-solvable in the differential case and 1-solvable in the difference case, the existence of hyperexponential solutions is equivalent to the existence of right factors of degree 1 in both of those cases. We let $\theta : k \longrightarrow k$ be σ if $\delta = 0$, δ otherwise. Note that θ is k-pseudo linear in both cases. We proceed by reducing the problem of factoring in $k[x;\sigma,\delta]$ to finding hypexponential solutions of elements of $k[x;\sigma,\delta]$. There are algorithms for finding such solutions in the differential case when k is a Liouvillian extension of $\mathrm{Const}_\delta(k)$ [18], and in the difference case when $k = C(t)$ for a subfield C where σ is the identity on C, t is transcendental over C and $\sigma t = t + 1$ [14]. Thus the algorithm presented here is complete for those fields. For more general coefficient fields or automorphisms (for example for q-difference operators), the discovery of an algorithm for computing hyperexponential solutions would yield a factoring algorithm for the corresponding skew polynomials.

The basic idea behind the factoring algorithm is trial division: let

$$p = x^n - \sum_{i=0}^{n-1} a_i x^i \quad \in k[x;\sigma,\delta] \tag{19}$$

and suppose that

$$q = x^m - \sum_{i=0}^{m-1} b_i x^i \tag{20}$$

is a right factor of p. If we can determine the b_i's up to some undetermined constants c_{ij}, then equating the right-remainder of p by q to 0 yields a system of algebraic equations with coefficients in $C = \mathrm{Const}_{\sigma,\delta}(k)$ for the c_{ij}. If that system has no solution in C then p does not have a right factor of degree m in $k[x;\sigma,\delta]$, otherwise any solution gives rises to such a factor (the same applies if C is replaced by an algebraic extension of C). Therefore, we proceed to determine the b_i's. Since $\sigma = 1_k$ or $\delta = 0_k$, we know from the theory of linear differential and difference equations that there exists a field extension K of k such that θ can be extended to either an automorphism or derivation of K, and $y_1, \ldots, y_m \in K$, linearly independent over $\mathrm{Const}_{\sigma,\delta}(K)$, such that $q *_\theta y_j = 0$ for each j. Note then that each y_j is a zero of p by Lemma 4. Since K is an algebra over k, Theorem 8 reduces the problem of determining the b_i's to determining $[\bar{i}]$ for $1 \leqslant i \leqslant m+1$. Since the nonzero $[\bar{i}]$'s are hyperexponential over k by Lemma 6, it is sufficient to compute an annihilator for each $[\bar{i}]$ and use the given algorithm for computing hyperexponential solutions in order to get candidates for the $[\bar{i}]$'s, thus completing the algorithm. Such annihilators are called the *associated operators of* p, and we describe in the rest of this section how to compute them from the coefficients of p.

4.3. Some elementary set operations

Let $n \geqslant m > 0$ be integers, and $S = \{s_1, \ldots, s_m\}$ a set of m integers with $1 \leqslant s_1 < \cdots < s_m \leqslant n$. We describe some basic operations on S which are needed by the algorithm. The first operation is *increment the elements of* S, and we denote the result by S^+, i.e.

$$S^+ = \{s \mid s-1 \in S\}.$$

The second operation is *increment the* kth *element of* S, and we denote the result by S_k^+, i.e.

$$S_k^+ = (S \cup \{1 + s_k\}) \setminus \{s_k\}.$$

The third operation is *replace the* kth *element of* S *by* l *and sort the result*, and we denote the result by $S_k^{[l]}$, i.e.

$$S_k^{[l]} = (S \cup \{l\}) \setminus \{s_k\} \quad \text{(sorted)}.$$

Finally, we define $\delta_k^{[l]}(S)$ to be

$$\delta_k^{[l]}(S) = \#\{s \in S \text{ such that } l < s < s_k\}$$

i.e. the number of elements of S which are strictly in between l and s_k.

We can apply the above set operations to the minors of a rectangular matrix: let R be any commutative ring, and $\mathcal{M}$ an $n \times m$ matrix with coefficients in R. For any set S as above, we define

$$[S]_k^+ = \begin{cases} [S_k^+] & \text{if } 1 + s_k \notin S \cup \{n+1\} \\ 0 & \text{if } 1 + s_k \in S \cup \{n+1\} \end{cases}$$

and for $1 \leqslant l \leqslant n$,

$$[S]_k^{[l]} = \begin{cases} (-1)^{\delta_k^{[l]}(S)}[S_k^{[l]}] & \text{if } l \notin S \setminus \{s_k\} \\ 0 & \text{if } l \in S \setminus \{s_k\} \end{cases}$$

4.4. The associated linear system

Let $n \geqslant m > 0$ be integers, $p \in k[x; \sigma, \delta]$ be given by (19), $y_1, \ldots, y_m$ be zeros of p in some compatible extension of k, and $\mathcal{M}$ their nth generalized Wronskian given by (17). We first need straightforward generalizations of Lemmas 4.1.4 and 4.1.5 of [21].

Lemma 7. *Let $S = \{s_1, \ldots, s_m\}$ be a set of m integers with $1 \leqslant s_1 < \cdots < s_m < n$. Then,*

$$\theta[S] = \begin{cases} \sum_{k=1}^m [S]_k^+ & \text{in the differential case,} \\ [S^+] & \text{in the difference case.} \end{cases}$$

Proof. Let S_m be the permutation group on m elements, and write $S = \{s_1, \ldots, s_m\}$ where $1 \leqslant s_1 < s_2 < \cdots < s_m < n$. Then, $[S] = \det(a_{i,j})$ where $a_{i,j} = \theta^{s_i - 1} y_j$, so

$$\begin{aligned} \theta[S] &= \theta\left(\sum_{\sigma \in S_m} (-1)^\sigma a_{1,\sigma(1)} \ldots a_{m,\sigma(m)}\right) \\ &= \sum_{k=1}^m \sum_{\sigma \in S_m} (-1)^\sigma a_{1,\sigma(1)} \ldots a_{k-1,\sigma(k-1)}\, \theta\left(a_{k,\sigma(k)}\right)\, a_{k+1,\sigma(k+1)} \ldots a_{m,\sigma(m)} \\ &= \sum_{k=1}^m [S]_k^+ \end{aligned}$$

in the differential case, and

$$\theta[S] = \theta\left(\sum_{\sigma \in S_m} (-1)^\sigma a_{1,\sigma(1)} \ldots a_{m,\sigma(m)}\right) = \sum_{\sigma \in S_m} (-1)^\sigma \theta a_{1,\sigma(1)} \ldots \theta a_{m,\sigma(m)} = [S^+]$$

in the difference case. □

Lemma 8. *Let $S = \{s_1, \ldots, s_m\}$ be a set of m integers with $1 \leqslant s_1 < \cdots < s_m = n$. Then,*

$$\theta[S] = \begin{cases} \sum_{k=1}^{m-1} [S]_k^+ + \sum_{j=0}^{n-1} a_j [S]_m^{[j+1]} & \textit{in the differential case,} \\ \sum_{j=0}^{n-1} a_j [S^+]_m^{[j+1]} & \textit{in the difference case.} \end{cases}$$

Proof. Let S_m be the permutation group on m elements, and write $S = \{s_1, \ldots, s_m\}$ where $1 \leqslant s_1 < s_2 < \cdots < s_m = n$. Then, $[S] = \det(a_{i,j})$ where $a_{i,j} = \theta^{s_i - 1} y_j$. In the differential case,

$$\theta[S] = \sum_{k=1}^{m-1} [S]_k^+ + \det(A)$$

where the first $m-1$ rows of A are rows s_1 to s_{m-1} of $\mathscr{M}$, and its last row is

$$(\theta a_{m1}, \ldots, \theta a_{mm}) = (\theta^n y_1, \ldots, \theta^n y_m) = \left(\sum_{j=0}^{n-1} a_j \theta^j y_1, \ldots, \sum_{j=0}^{n-1} a_j \theta^j y_m \right).$$

Hence, $\det(A) = \sum_{j=0}^{n-1} a_j \det(A_j)$ where the first $m-1$ rows of A_j are rows s_1 to s_{m-1} of $\mathscr{M}$, and the last row of A_j is $\left(\theta^j y_1, \ldots, \theta^j y_m\right)$, i.e. the $(j+1)$th row of $\mathscr{M}$. If $j+1 \in S \setminus \{n\}$, then this row appears twice in A_j, so $\det(A_j) = 0$. Otherwise, it takes $\delta_m^{[j+1]}(S)$ row exchanges to turn A_j into $\mathscr{M}_{S_m^{[j+1]}}$, so $\det(A_j) = (-1)^{\delta_m^{[j+1]}(S)}[S_m^{[j+1]}]$. Hence,

$$\theta[S] = \sum_{k=1}^{m-1} [S_k^+] + \sum_{j=0}^{n-1} a_j \det(A_j) = \sum_{k=1}^{m-1} [S]_k^+ + \sum_{j=0}^{n-1} a_j [S]_m^{[j+1]}.$$

In the difference case,

$$\theta[S] = \det(A)$$

where the first $m-1$ rows of A are rows $s_1 + 1$ to $s_{m-1} + 1$ of $\mathscr{M}$, and its last row is

$$(\theta a_{m1}, \ldots, \theta a_{mm}) = (\theta^n y_1, \ldots, \theta^n y_m) = \left(\sum_{j=0}^{n-1} a_j \theta^j y_1, \ldots, \sum_{j=0}^{n-1} a_j \theta^j y_m \right).$$

Hence, $\det(A) = \sum_{j=0}^{n-1} a_j \det(A_j)$ where the first $m-1$ rows of A_j are rows $s_1 + 1$ to $s_{m-1} + 1$ of $\mathscr{M}$, and the last row of A_j is $(\theta^j y_1, \ldots, \theta^j y_m)$, i.e. the $(j+1)$th row of $\mathscr{M}$. If $j+1 \in S^+ \setminus \{n+1\}$, then this row appears twice in A_j, so $\det(A_j) = 0$. Otherwise, it takes $\delta_m^{[j+1]}(S^+)$ row exchanges to turn A_j into $\mathscr{M}_{(S^+)_m^{[j+1]}}$, so $\det(A_j) = (-1)^{\delta_m^{[j+1]}(S^+)}[(S^+)_m^{[j+1]}]$. Hence $\det(A_j) = [S^+]_m^{[j+1]}$ and the lemma follows. □

Let w be the vector composed of all the N minors of size m of $\mathscr{M}$, where $N = \binom{n}{m}$. As a consequence of Lemmas 7 and 8, the linear subspace generated by w over k is closed under θ, i.e. there exists an $N \times N$ matrix $M_m(p)$ with coefficients in k such that

$$\theta w = M_m(p) \cdot w . \tag{21}$$

The above system is called the mth *associated system of p*. In order to compute annihilators for any $[S]$, we need to uncouple the above system. An algorithm for doing this is described in the next section.

Example. Let $L = D^4 - a_3D^3 - a_2D^2 - a_1D - a_0$ be the generic differential operator of order 4, and let us compute its 2nd associated system.

1. The 2 by 2 minors are: $[1,2],[1,3],[2,3],[1,4],[2,4],[3,4]$.
2. Applying Lemma 7 to those subsets which do not contain 4 we get:

$$\begin{aligned}
\theta[1,2] &= [1,2]_1^+ + [1,2]_2^+ = 0 + [\{1,2\}_2^+] = [1,3]\\
\theta[1,3] &= [1,3]_1^+ + [1,3]_2^+ = [\{1,3\}_1^+] + [\{1,3\}_2^+] = [2,3] + [1,4]\\
\theta[2,3] &= [2,3]_1^+ + [2,3]_2^+ = 0 + [\{2,3\}_2^+] = [2,4] .
\end{aligned}$$

Furthermore, applying Lemma 8 to those subsets which do contain 4 we get:

$$\begin{aligned}
\theta[1,4] &= [1,4]_1^+ + a_3\,[1,4]_2^{[4]} + a_2\,[1,4]_2^{[3]} + a_1\,[1,4]_2^{[2]} + a_0[1,4]_2^{[1]}\\
&= [\{1,4\}_1^+] + a_3\,[\{1,4\}_2^{[4]}] + a_2\,[\{1,4\}_2^{[3]}] + a_1\,[\{1,4\}_2^{[2]}] + 0\\
&= [2,4] + a_3\,[1,4] + a_2\,[1,3] + a_1\,[1,2]\\
\theta[2,4] &= [2,4]_1^+ + a_3\,[2,4]_2^{[4]} + a_2\,[2,4]_2^{[3]} + a_1\,[2,4]_2^{[2]} + a_0[2,4]_2^{[1]}\\
&= [\{2,4\}_1^+] + a_3\,[\{2,4\}_2^{[4]}] + a_2\,[\{2,4\}_2^{[3]}] + 0 - a_0\,[\{2,4\}_2^{[1]}]\\
&= [3,4] + a_3\,[2,4] + a_2\,[2,3] - a_0\,[1,2]\\
\theta[3,4] &= [3,4]_1^+ + a_3\,[3,4]_2^{[4]} + a_2\,[3,4]_2^{[3]} + a_1\,[3,4]_2^{[2]} + a_0[3,4]_2^{[1]}\\
&= 0 + a_3\,[\{3,4\}_2^{[4]}] + 0 - a_1\,[\{3,4\}_2^{[2]}] - a_0\,[\{3,4\}_2^{[1]}]\\
&= a_3\,[3,4] - a_1\,[2,3] - a_0\,[1,3]
\end{aligned}$$

3. Hence, the generic 2nd associated system for differential operators of order 4 is:

$$\theta \begin{bmatrix} [1,2]\\ [1,3]\\ [2,3]\\ [1,4]\\ [2,4]\\ [3,4] \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ a_1 & a_2 & 0 & a_3 & 1 & 0\\ -a_0 & 0 & a_2 & 0 & a_3 & 1\\ 0 & -a_0 & -a_1 & 0 & 0 & a_3 \end{bmatrix} \begin{bmatrix} [1,2]\\ [1,3]\\ [2,3]\\ [1,4]\\ [2,4]\\ [3,4] \end{bmatrix} \tag{22}$$

Example. Let $L = E^4 - a_3E^3 - a_2E^2 - a_1E - a_0$ be the generic difference operator of order 4. The 2nd associated system for L is:

$$\theta \begin{bmatrix} [1,2] \\ [1,3] \\ [2,3] \\ [1,4] \\ [2,4] \\ [3,4] \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -a_0 & 0 & a_2 & 0 & a_3 & 0 \\ 0 & -a_0 & -a_1 & 0 & 0 & a_3 \\ 0 & 0 & 0 & -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} [1,2] \\ [1,3] \\ [2,3] \\ [1,4] \\ [2,4] \\ [3,4] \end{bmatrix} \qquad (23)$$

4.5. *The associated operators*

We describe in this section a method for uncoupling the system (21) which yields annihilators for any $[S]$. Note that any other uncoupling method, for example cyclic vectors or block-diagonal decomposition [22] could also be used. We first define the sequence of matrices $M_{m,1}(L), \ldots, M_{m,N}(L)$, where $N = \binom{n}{m}$, by $M_{m,1}(L) = M_m(L)$ and, for $2 \leqslant i \leqslant N$,

$$M_{m,i}(L) = \begin{cases} M_{m,i-1}(L)\, M_m(L) + \theta M_{m,i-1}(L) & \text{in the differential case,} \\ \theta M_{m,i-1}(L)\, M_m(L) & \text{in the difference case.} \end{cases} \qquad (24)$$

where θ is applied componentwise to matrices. It can be easily checked by induction that

$$\theta^i w = M_{m,i}(L)\, w \quad \text{for each } i\,.$$

For any subset S of m integers in $\{1, \ldots, n\}$, we write w_S^* for the vector $[\theta w_S, \theta^2 w_S, \ldots, \theta^N w_S]^{\mathrm{T}}$ and n_S for the index of S in our ordering of those subsets. Define A_S to be the $N \times N$ matrix such that the ith row of A_S is the n_Sth row of $M_{m,i}$ for each i. We then have

$$A_S\, w = w_S^*\,.$$

Nondegenerate case: A_S is nonsingular for some S. In this case, the equation with $[S]$ in the left hand side in the system $w = A_S^{-1} w_S^*$ is an annihilator for $[S]$, while all the other equations give formulas for all the other $[T]$'s as linear combinations of $\theta[S], \ldots, \theta^N[S]$. If the annihilator for $[S]$ has no hyperexponential solution, then $[S] = 0$, so $[1, \ldots, m] = 0$ (since it is a linear combination of $\theta[S], \ldots, \theta^N[S]$), which implies that p has no right factor of degree m. Otherwise, the expressions for $[S]$ with undetermined constants yield all the possible candidate right factors of p of degree m.

Degenerate case: A_S is singular for all S. In that case, let $(u_1, \ldots, u_q)$ be a basis for the kernel of the transpose of A_S for some S. Since each $u_i \in k^N$ corresponds to a linear dependence for the rows of A_S, each dot product $u_i \cdot [S^*]$ gives an annihilator for $[S]$. If $q = N - \mathrm{rank}(A_S) > 1$, we obtain an overdetermined system of associated equations for $[S]$, which can be reduced to one equation by taking their right gcd. We first compute the annihilator for $[1,\dots,m]$ in this case. If it has no hyperexponential solution, then $[1,\dots,m]=0$ which implies that p has no right factor of degree m. Otherwise, using Gaussian elimination on $A=A_{\{1,\dots,m\}}$, we get an invertible matrix B and an upper triangular matrix U such that $A=BU$, so it may be possible to obtain expressions for some other $[\bar{i}]$'s as linear combinations of $\theta[1,\dots,m],\dots,\theta^N[1,\dots,m]$ from the equations $Uw=B^{-1}[1,\dots,m]^*$. We then generate the associated equation(s) for the next $[\bar{i}]$ which we need to compute, and either look for its hyperexponential solutions over k, or replace $[\bar{i}]$ by $(-1)^{m-i}[1,\dots,m]b_i$ and look for all the solutions b_i in k [1,4]. We repeat this process until candidates for all the $[\bar{i}]$'s are found, noting that after each step, a decompostion $A=BU$ may yield expressions for some other $[\bar{j}]$'s.

We should mention at this point a further improvement of Tsarev [20]: the $[S]$'s must satisfy the Grassmann–Plücker relations [19], so evaluating those relations on the candidates, we have, for all the $[S]$'s, yields algebraic equations on the undetermined parameters. In both the degenerate and nondegenerate cases, those conditions are necessary in order for the candidates to correspond to an actual factor of the initial operator. In the case of differential operators, Tsarev also states that those are necessary and sufficient conditions in the nondegenerate case, a fact which can be used to avoid the final trial division altogether.

Example. Consider the operator

$$L=D^4-2tD^2-2D+t^2\in\mathbb{C}(t)[D;1,\mathrm{d}/\mathrm{d}t]$$

which has no hyperexponential solution, hence no right factor of degree 1.

Its associated system is $\theta w=M_2w$ where M_2 is obtained from (22) by substitution:

$$M_2=\begin{bmatrix}0&1&0&0&0&0\\0&0&1&1&0&0\\0&0&0&0&1&0\\2&2t&0&0&1&0\\t^2&0&2t&0&0&1\\0&t^2&-2&0&0&0\end{bmatrix}$$

There is only one nonsingular A_S, namely $A_{[3,4]}$ whose inverse yields an annihilator of degree 6 for $[3,4]$. This annihilator has a linear space of hyperexponential solutions generated by $\{t^3-2,t^2,t\}$, so if there is a right factor of degree 2 we must have

$$[3,4]=a(t^3-2)+bt^2+ct$$

for some $a,b,c\in\mathbb{C}$. Plugging this form back into the expressions for the other minors given by $A^{-1}_{[3,4]}$ we get

$$[1,2]=at+b,\quad [1,3]=a,\quad [2,3]=-at^2-bt-\frac{c}{2},\quad [1,4]=at^2+bt+\frac{c}{2},$$
$$[2,4]=-2at-b.$$

The (4,2) Grassmann–Plücker relation are generated by [19]:

$$[1,2][3,4] - [1,3][2,4] + [1,4][2,3] = 0 \tag{25}$$

so, replacing the various $[i,j]'s$ by the above values, we get $-ab - c^2/4 = 0$ which means that either $a \neq 0$ or $b \neq 0$, and that $a = -c^2/(4b)$. Thus, if $D^2 + b_1D + b_0$ is a right factor of L, we must have

$$b_1 = -\frac{[1,3]}{[1,2]} = -\frac{a}{at+b} = -\frac{c^2}{c^2t - 4b^2}$$

and

$$b_0 = \frac{[2,3]}{[1,2]} = \frac{-at^2 - bt - c/2}{at+b} = -\frac{c^2t^2 - 4b^2t - 2bc}{c^2t - 4b^2}.$$

Dividing L by $D^2 + b_1D + b_0$ on the right gives a remainder of 0, hence, for any constants b and c not both 0:

$$\begin{aligned} & D^4 - 2tD^2 - 2D + t^2 \\ & = \left(D^2 + \frac{c^2}{c^2t - 4b^2}D - \frac{c^4t^3 - 8b^2c^2t^2 + 2b(c^3 + 8b^3)t + c^4 - 8b^3c}{(c^2t - 4b^2)^2}\right) \\ & \times \left(D^2 - \frac{c^2}{c^2t - 4b^2}D - \frac{c^2t^2 - 4b^2t - 2bc}{c^2t - 4b^2}\right) \end{aligned}$$

For $c = 0$ we get $D^4 - 2tD^2 - 2D + t^2 = (D^2 - t)(D^2 - t)$. If we do not check the relation (25) before doing the trial division, equating the remainder to 0 would have also yielded the condition $c^2 + 4ab = 0$.

5. A normal form for pseudo-linear maps

Solving a system of linear equations of the form (2,4) is equivalent to solving an equation of the form $\theta x = 0$ or $\theta x = v$ where θ is some pseudo-linear operator in a finite dimensional vector space V. For example, the differential system

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{x^2}\begin{bmatrix} x & -x^2 - 1 \\ x^4 + 2x^2 & -2x^3 - x \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{26}$$

is equivalent to $\theta y = 0$ where θ is the operator of the last example of Section 1.3. As seen there, the change of basis

$$\begin{cases} z_1 = -xy_1 + y_2 \\ z_2 = 2y_1 - y_2/x \end{cases} \tag{27}$$

yields the uncoupled system

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = -\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

which can be solved by trivial methods, yielding the solutions

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \left(c_1\begin{bmatrix} 1 \\ -x \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)\mathrm{e}^{-x^2/2}.$$

The solutions (y_1, y_2) of (26) can then be found by inverting (27).

Let k, σ, δ be as previously and V a finite dimensional vector space of dimension N over k. A natural question is then to ask whether for any pseudo-linear operator, there exists a basis of V in which the matrix of θ has a special form, for example diagonal, triangular, or companion. B. Zürcher [22] has recently generalized Danilevski's weak Frobenius algorithm [9] to show that any pseudo-linear map can be brought via a change of basis to a block-diagonal form, where each block is a companion matrix of the form

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & 0 & 1 \\ c_0 & c_1 & \dots & c_{M-2} & c_{M-1} \end{bmatrix} \tag{28}$$

Theorem 10 (Zürcher [22]). *For any pseudo-linear map $\theta : V \to V$, we can compute a basis for V over k such that the matrix M of θ w.r.t. to this basis is of the form $M = diag(C_1, C_2, \ldots, C_m)$ where the C_i's are all companion matrices.*

We now briefly outline his algorithm, which proceeds by successive change of basis by elementary matrices, corresponding to elementary row and column operations (see [22] for details and proofs of correctness). The first elementary matrix is $D_i(a)$ for $a \in k$, which is the identity matrix with an a in the (i,i)th entry. By the change of basis formula (11) the effect of the change of basis $\mathscr{E} = D_i(a)\mathscr{B}$ on $M_{\mathscr{B}}(\theta)$ is

- multiply column i by $\sigma(a)$,
- multiply row i by a^{-1},
- add $a^{-1}\delta a$ to the (i,i)th entry, in that order. The next elementary matrix is $C_{ij}(a)$ for $a \in k$ and $i \neq j$, which is the identity matrix with an a in the (i,j)th entry. The effect of the change of basis $\mathscr{E} = C_{ij}(a)\mathscr{B}$ on $M_{\mathscr{B}}(\theta)$ is
- add $\sigma(a)$ times column i to column j,

- add $-a$ times row j to row i,
- add δa to the (i,j)th entry in that order. The last elementary matrix is the usual P_{ij} which is the identity matrix with rows i and j exchanged. Its action is to exchange columns i and j, and rows i and j in $M_{\mathscr{B}}(\theta)$.

The algorithm proceeds recursively by increasing the size of a companion block or moving to the next block. By taking $i = 1$ if necessary, we can assume that the matrix of θ has the form

$$\begin{array}{c|cccccccc|}
 & & & & \overset{i}{\downarrow} & & & & \\
 & 0 & 1 & & 0 & 0 & \dots & \dots & 0 \\
 & \vdots & & \ddots & & \vdots & & & \vdots \\
 & 0 & \dots & & 1 & 0 & \dots & \dots & 0 \\
i\rightarrow & * & \dots & \dots & * & \circledast & \dots & \dots & \circledast \\
 & * & \dots & \dots & * & * & \dots & \dots & * \\
 & \vdots & & & \vdots & \vdots & & & \vdots \\
 & \vdots & & & \vdots & \vdots & & & \vdots \\
 & * & \dots & \dots & * & * & \dots & \dots & *
\end{array} \tag{29}$$

for some $i < N$. Suppose that $m_{i,j} \neq 0$ for some j, $i < j \leqslant N$, i.e. one of the $\circledast$'s is nonzero. Applying $P_{i+1,j}$ we get a matrix of the form (29) with $m_{i,i+1} \neq 0$. Applying then $D_{i+1}\left(\sigma^{-1}\left(m_{i,i+1}^{-1}\right)\right)$ replaces $m_{i,i+1}$ by 1 without modifying the first $i-1$ rows. With $m_{i,i+1} = 1$, applying $C_{i+1,j}\left(\sigma^{-1}(-m_{i,j})\right)$ for $j \neq i+1$ replaces $m_{i,j}$ by 0 without modifying the first $i-1$ rows, so doing this for $j = 1,\dots,i,i+2,\dots,N$ yields a matrix of the form (29) with i replaced by $i+1$, i.e. with a larger companion block. We can repeat this procedure until we get a matrix of the form

$$\begin{array}{c|cccccccc|}
 & & & & \overset{i}{\downarrow} & & & & \\
 & 0 & 1 & & 0 & 0 & \dots & \dots & 0 \\
 & \vdots & & \ddots & & \vdots & & & \vdots \\
 & 0 & \dots & & 1 & 0 & \dots & \dots & 0 \\
i\rightarrow & * & \dots & \dots & * & 0 & \dots & \dots & 0 \\
 & * & \dots & \dots & * & * & \dots & \dots & * \\
 & \vdots & & & \vdots & \vdots & & & \vdots \\
 & \vdots & & & \vdots & \vdots & & & \vdots \\
 & * & \dots & \dots & * & * & \dots & \dots & *
\end{array} \tag{30}$$

for some $i < N$. At this point, applying $C_{j,i-1}(m_{j,i})$ for $j = i+1$ replaces $m_{j,i}$ by 0 without modifying the companion block already obtained or the zeros to its right, so doing this for $j = i+1,\ldots,N$ sets the ith column under the companion block to 0. We repeat this process for columns $i-1$ down to 2, obtaining a matrix of the form

$$\begin{array}{cc}
 & \begin{array}{cccccccc} & & & i & & & & \\ & & & \downarrow & & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i\rightarrow \\ \\ \\ \\ \\ \end{array} &
\left(\begin{array}{cccccccc}
0 & 1 & & 0 & 0 & \ldots & \ldots & 0 \\
\vdots & & \ddots & & \vdots & & & \vdots \\
0 & \ldots & & 1 & 0 & \ldots & \ldots & 0 \\
* & \ldots & \ldots & * & 0 & \ldots & \ldots & 0 \\
\circledast & 0 & \ldots & 0 & * & \ldots & \ldots & * \\
\vdots & \vdots & & \vdots & \vdots & & & \vdots \\
\vdots & \vdots & & \vdots & \vdots & & & \vdots \\
\circledast & 0 & \ldots & 0 & * & \ldots & \ldots & *
\end{array}\right)
\end{array} \qquad (31)$$

Suppose that $m_{j,1} = 0$ for all j, $i < j \leqslant N$, i.e. all the $\circledast$'s are zero. Then, (31) is a block-diagonal matrix with an $i \times i$ companion matrix of the form (28) followed by an $(N-i) \times (N-i)$ square block, to which we can apply this algorithm recursively, completing the decomposition.

So we can assume that $m_{j,1} \neq 0$ for some j, $i < j \leqslant N$. Applying $P_{j,N}$ we get a matrix of the form (31) with $m_{N,1} \neq 0$. Applying then $D_N(m_{N,1})$ replaces $m_{N,1}$ by 1 without changing the shape (31) of the matrix. With $m_{N,1} = 1$, applying $C_{j,N}(m_{j,N})$ for $i < j < N$ replaces each $m_{j,1}$ by 0, so we obtain a matrix of the form

$$\begin{array}{cc}
 & \begin{array}{cccccccc} & & & i & & & & \\ & & & \downarrow & & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i\rightarrow \\ \\ \\ \\ \\ \end{array} &
\left(\begin{array}{cccc|ccc|c}
0 & 1 & & 0 & 0 & \ldots & 0 & 0 \\
\vdots & & \ddots & & \vdots & & \vdots & \vdots \\
0 & \ldots & & 1 & \vdots & & \vdots & \vdots \\
* & \ldots & \ldots & * & 0 & \ldots & 0 & 0 \\
\hline
0 & \ldots & \ldots & 0 & * & \ldots & * & * \\
\vdots & & & \vdots & \vdots & & \vdots & \vdots \\
0 & \ldots & \ldots & 0 & * & \ldots & * & * \\
\hline
1 & 0 & \ldots & 0 & * & \ldots & * & *
\end{array}\right)
\end{array} \qquad (32)$$

Applying all the permutations $P_{j,N}$ for $j = 1,\dots,N-1$ in sequence performs a rotation of the columns towards the right and of the rows towards the bottom, transforming (32) into a matrix of the form

$$\begin{array}{r|c|cccc|ccc|}
 & & & & & i+1 & & & \\
 & & & & & \downarrow & & & \\
\cline{2-9}
 & * & 1 & 0 & \cdots & 0 & * & \cdots & * \\
\cline{2-9}
 & 0 & 0 & 1 & & 0 & 0 & \cdots & 0 \\
 & \vdots & \vdots & & \ddots & & \vdots & & \vdots \\
 & \vdots & 0 & \cdots & & 1 & \vdots & & \vdots \\
i+1\rightarrow & 0 & * & \cdots & \cdots & * & 0 & \cdots & 0 \\
\cline{2-9}
 & * & 0 & \cdots & \cdots & 0 & * & \cdots & * \\
 & \vdots & \vdots & & & \vdots & \vdots & & \vdots \\
 & * & 0 & \cdots & \cdots & 0 & * & \cdots & * \\
\cline{2-9}
\end{array} \qquad (33)$$

Applying $C_{2,1}\left(\sigma^{-1}(-m_{1,1})\right)$ replaces $m_{1,1}$ by 0 but modifies $m_{2,1}, m_{2,2}$ and $m_{2,i+2}$ to $m_{2,N}$. Applying $C_{3,j}\left(\sigma^{-1}(-m2,j)\right)$ for $j = 1, 2, i+2, \dots, N$ replaces $m_{2,1}, m_{2,2}$ and $m_{2,i+2}$ to $m_{2,N}$ by 0, but modifies $m_{3,1}$ to $m_{3,3}$ and $m_{3,i+2}$ to $m_{3,N}$. Doing this for rows 1 to i yields a matrix of the form (33) but with $m_{1,1} = 0$ and row $i+1$ arbitrary. This is then a matrix of the form (29) with i replaced by $i+1$, i.e. with a larger companion block, so we can repeat this algorithm until complete decomposition is obtained.

Zürcher's algorithm can be used in general to reduce systems of equations of the form $\theta x = v$ to higher order uncoupled equations, which can then be passed to the skew-polynomial factorisation algorithm. In the remainder of this section, we illustrate this process via examples of systems of linear differential and difference equations.

5.1. *Differential equations*

Let

$$y' = My + v \qquad (34)$$

be a first order differential system where M is an $m \times m$ matrix with entries in a differential field k with derivation $'$, and $v \in k^m$. Let K be any differential extension of k, $\sigma = 1_K$, $\delta : K \to K$ be given by $\delta a = -a'$, and $\theta : K^m \to K^m$ be the pseudo-linear map whose matrix w.r.t. the canonical basis is M. Zürcher's algorithm produces an invertible matrix A and companion matrices $C_1, \dots, C_q$ such that for $z = A^{-1}y$, $\theta z = \operatorname{diag}(C_1, C_2, \dots, C_q) \cdot z - z'$. Assuming without loss of generality that there is only one companion block C of the form (28), we get that $y \in K^m$ is a solution of (34) if and only if $z = A^{-1}y$ is a solution of

$$z' = Cz + w \quad \text{where } w = A^{-1}v \in k^m. \tag{35}$$

System (35) is of the form

$$z_{i+1} = z_i' - w_i = z_1^{(i)} - \sum_{j=1}^{i} w_j^{(i-j)} \quad \text{for } 1 < i < m$$

and the last equation is $z_m' = \sum_{i=0}^{m-1} c_i z_{i+1} + w_m$ which becomes

$$z_1^{(m)} - \sum_{j=1}^{m} w_j^{(m-j)} = \sum_{i=0}^{m-1} c_i z_1^{(i)} - \sum_{i=0}^{m-1} c_i \sum_{j=1}^{i} w_j^{(i-j)} + w_m \tag{36}$$

which is an uncoupled differential equation for z_1.

Example. (*Barkatou* [3]). Consider the differential system

$$t^2\frac{\mathrm{d}y}{\mathrm{d}t} = My = \begin{bmatrix} 4t+1 & -5t & 7t & -8t & 8t & -6t \\ -10t & 9t+1 & -14t & 16t & -16t & 12t \\ -5t & 5t & -8t+1 & 8t & -8t & 6t \\ 10t & -10t & 14t & -17t+1 & 16t & -12t \\ 5t & -5t & 7t & -8t & 7t+1 & -6t \\ -5t & 5t & -7t & 8t & -8t & 5t+1 \end{bmatrix} \cdot y$$

Taking $' = t^2\mathrm{d}/\mathrm{d}t$, $\sigma = 1$, $\delta = -t^2\mathrm{d}/\mathrm{d}t$ and applying Zürcher's algorithm to the pseudo-linear map whose matrix is M, we get the invertible matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{-5t+1}{5t} & -\frac{1}{5t} & \frac{7}{5} & -\frac{8}{5} & \frac{8}{5} & -\frac{6}{5} \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the companion matrices

$$C_1 = \begin{bmatrix} 0 & 1 \\ 5t^2 - 5t - 1 & 5t + 2 \end{bmatrix}, \quad C_2 = C_3 = C_4 = C_5 = [1 - t]$$

such that y is a solution of $t^2\mathrm{d}y/\mathrm{d}t = My$ if and only if $z = A^{-1}y$ is a solution of $t^2\mathrm{d}z/\mathrm{d}t = Tz$ where $T = \mathrm{diag}(C_1, C_2, C_3, C_4, C_5)$. Using (36) we get the uncoupled equations $z_2 = t^2\mathrm{d}z_1/\mathrm{d}t$, $t^2\mathrm{d}z_i/\mathrm{d}t = (1-t)z_i$ for $i \in \{3, 4, 5, 6\}$, and

$$t^2\frac{\mathrm{d}}{\mathrm{d}t}\left(t^2\frac{\mathrm{d}z_1}{\mathrm{d}t}\right) = \left(5t^2 - 5t - 1\right)z_1 + (5t+2)t^2\frac{\mathrm{d}z_1}{\mathrm{d}t}.$$

The corresponding skew polynomial can be factored as a product of operators of degree 1 [18], yielding the general solution

$$z_1 = \left(c_1 t^5 + \frac{c_2}{t}\right)\mathrm{e}^{-1/t}$$

while for $i \in \{3,4,5,6\}$, the equations $t^2 \mathrm{d}z_i/\mathrm{d}t = (1-t)z_i$ have general solutions $z_i = c_i \mathrm{e}^{-1/t}/t$. Using $y = Az$ yields the general solution of the original system.

5.2. *Difference equations*

Let

$$Ey = My + v \tag{37}$$

be a first order recurrence system where M is an $m \times m$ matrix with entries in a difference field k with transform E (for example, which sends n to $n+1$), and $v \in k^m$. Let K be any difference extension of k, $\sigma = E^{-1}$, $\delta = \sigma - 1_K$, and $\theta : K^m \to K^m$ be the pseudo-linear map whose matrix w.r.t. the canonical basis is $\sigma(M) - I$. Zürcher's algorithm produces an invertible matrix A and companion matrices $C_1, \ldots, C_q$ such that for $z = A^{-1}y$, $\theta z = \mathrm{diag}(C_1, C_2, \ldots, C_q) \cdot \sigma(z) + \delta z$. Assuming without loss of generality that there is only one companion block C of the form (28), we get that $y \in K^m$ is a solution of (37) if and only if $z = A^{-1}y$ is a solution of

$$C\sigma(z) + \delta z = A^{-1}v\,.$$

Applying E on both sides of that system, we get $E(C)z + z - Ez = w$ where $w = E(A^{-1}v)$, hence

$$Ez = Tz - w \quad \text{where } T = E(C) + I\,. \tag{38}$$

System (38) is of the form

$$z_{i+1} = Ez_i - z_i + w_i = \Delta^i z_1 + \sum_{j=1}^{i} E^{i-j} w_j \quad \text{for } 1 < i < m\,,$$

where $\Delta = E - 1_K$ is the associated difference operator. The last equation is $Ez_m = \sum_{i=0}^{m-1} E(c_i) z_{i+1} + z_m + w_m$ which becomes

$$\Delta^n z_1 + \sum_{j=1}^{i} \Delta E^{i-j} w_j = \sum_{i=0}^{m-1} E(c_i)\Delta^i z_1 + \sum_{i=0}^{m-1} E(c_i) \sum_{j=1}^{i} E^{i-j} w_j + w_m \tag{39}$$

which is an uncoupled difference equation for z_1.

Example. Consider the recurrence system $y(n+1) = M(n)y(n)$ where $M(n)$ is

$$\begin{bmatrix} n & 1 & 0 & n+3 \\ 4n^3 + 34n^2 + 71n - 7 & 4n^2 + 36n + 80 & 3n + 13 & 4n^3 + 48n^2 + 186n + 233 \\ -2n^2 - 10n & -2n - 10 & -2 & -2n^2 - 16n - 29 \\ -4n^2 - 18n + 2 & -4n - 20 & -3 - \frac{1}{n+4} & -4n^2 - 32n - 58 - \frac{1}{n+4} \end{bmatrix}$$

Taking $En = n+1$, $\sigma n = n-1$, $\delta = \sigma - 1$ and applying Zürcher's algorithm to the pseudo-linear map whose matrix is $M(n-1) - I$, we get the invertible matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 4n^2+27n+49 & 1 & -3n-9 & -n-3 \\ -2n-8 & 0 & 1 & 0 \\ -4n-16 & 0 & 3 & 1 \end{bmatrix}$$

and the companion matrices

$$C_1 = \begin{bmatrix} 0 & 1 \\ 2n-1 & n-2 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0 & 1 \\ \frac{-4}{n+3} & -\frac{2n+7}{n+3} \end{bmatrix}$$

such that y is a solution of $y(n+1) = M(n)y(n)$ if and only if $z = A^{-1}y$ is a solution of $z(n+1) = Tz(n)$ where $T = \operatorname{diag}(C_1(n+1), C_2(n+1)) + 1$. Using (39) we get the uncoupled equations $z_2 = \Delta z_1$, $z_4 = \Delta z_3$,

$$\Delta^2 z_1 = (2n+1)z_1 + (n-1)\Delta z_1, \quad \text{and} \quad \Delta^2 z_3 = -\frac{4}{n+4}z_3 - \frac{2n+9}{n+4}\Delta z_3 .$$

The corresponding skew polynomials can be factored as products of operators of degree 1 [14], yielding the general solutions

$$z_1 = n!\left(c_1 + c_2\sum_{i=0}^{n}\frac{(-1)^i}{i!}\right), \quad z_3 = \frac{c_3 + c_4(-1)^n(2n+3)}{(n+1)(n+2)}$$

and $y = Az$ yields the general solution of the original system.

References

[1] S.A. Abramov, Rational solutions of linear differential and difference equations with polynomial coefficients, *Comput. Math. Math. Phys.* **29** (1989) 7–12 (translated from *Zh. Vychisl. Mat. i Mat. Fiz.* **29** (1989) 1611–1620).

[2] S.A. Abramov, D'Alembertian solutions of nonhomogeneous linear ordinary differential and difference equations, manuscript, 1994.

[3] M.A. Barkatou, An algorithm for computing a companion block diagonal form for a system of linear differential equations, *Appl Algebra Engrg. Comm. Comput.* **4** (1993) 185–195.

[4] M. Bronstein, On solutions of linear ordinary differential equations in their coefficient field, *J. Symbol. Comput.* **13** (1992) 413–439.

[5] M. Bronstein, Linear ordinary differential equations: breaking through the order 2 barrier, in: P. Wang, ed. *Proc. ISSAC'92* (1992) 42–48.

[6] M. Bronstein and M. Petkovšek, On Ore rings, linear operators and factorisation, *Programmirovanie* **20** (1994) 27–45. Also Research Report 200, Informatik, ETH Zürich.

[7] P.M. Cohn, Rings with a weak algorithm, *Trans. Amer. Math. Soc.* **109** (1963) 332–356.

[8] P.M. Cohn, *Free Rings and their Relations* (Academic Press, New York, 1971).

[9] A. Danilewski, The numerical solution of the secular equation, *Mat. Sbornik* **2** (1937) 169–171. (in Russian).

[10] N. Jacobson, Pseudo-linear transformations, *Ann. Math.* **38** (1937) 484–507.

[11] D.E. Knuth, *The Art of Computer Programming*, Vol. 2 (Addison-Wesley, Reading, MA, 1981).

[12] E. Landau, Ein Satz über die Zerlegung homogener linearer Differentialausdrücke in irreducible Factoren, *J. Reine Angew. Math.* **124** (1902) 115–120.

[13] O. Ore, Theory of non-commutative polynomials, *Ann. Math.* **34** (1933) 480–508.

[14] M. Petkovšek, Hypergeometric solutions of linear difference equations with polynomial coefficients, *J. Symbol. Comput.* **14** (1992) 243–264.

[15] M. Petkovšek and B. Salvy, Finding all hypergeometric solutions of linear differential equations, in: M. Bronstein, ed. *Proc. ISSAC'93,* (ACM Press, New York, 1993) 27–33.

[16] L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen* (Teubner, Leipzig, 1895).

[17] F. Schwarz, A factorization algorithm for linear ordinary differential equations, in: G. Gonnet, ed. *Proc. ISSAC'89*, (ACM Press, New York, 1989), 17–25.

[18] M. Singer, Liouvillian solutions of linear differential equations with Liouvillian coefficients, *J. Symbol. Comput.* **11** (1991) 251–273.

[19] B. Sturmfels, *Algorithms in Invariant Theory* (Springer, Wien, 1993) Chapter 3.

[20] S.P. Tsarev, On some problems in the factorization of linear ordinary differential operators: new applications of old results, *Programmirovanie* **20** (1994) 45–48.

[21] K. Wolf, *Effiziente Algorithmen zur Lösung linearer Differentialgleichungsysteme und zur Faktorisierung linearer Differentialoperatoren über liouvillischen Körpern,* Doctoral dissertation, Mathematics, RFW Universität, Bonn, 1992.

[22] B. Zürcher, *Rationale Normalform von pseudo-linearen Abbildungen,* Diplomarbeit Bericht, Mathematik, ETH Zürich, 1994.