On Computing the Hermite Form of a Matrix of Differential Polynomials

Mark Giesbrecht and Myung Sub Kim

Cheriton School of Computer Science, University of Waterloo, Waterloo, Canada

Abstract. Given a matrix $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ over the ring of differential polynomials, we show how to compute the Hermite form H of A and a unimodular matrix U such that $UA = H$. The algorithm requires a polynomial number of operations in F in terms of n, $\deg_{\mathcal{D}} A$, $\deg_t A$. When $\mathsf{F} = \mathbb{Q}$ it require time polynomial in the bit-length of the rational coefficients as well.

1 Introduction

Canonical forms of matrices over principal ideal domains (such as $\mathbb{Z}$ or $\mathsf{F}[x]$, for a field F) have proven invaluable for both mathematical and computational purposes. One of the successes of computer algebra over the past three decades has been the development of fast algorithms for computing these canonical forms. These include triangular forms such as the Hermite form (Hermite, 1863), low degree forms like the Popov form (Popov, 1972), as well as the diagonal Smith form (Smith, 1861).

Canonical forms of matrices over non-commutative domains, especially rings of differential and difference operators, are also extremely useful. These have been examined at least since Dickson (1923), Wedderburn (1932), and Jacobson (1943). A typical domain under consideration is that of differential polynomials. For our purposes these are polynomials over a function field $\mathsf{F}(t)$ (where F is a field of characteristic zero, typically an extension of $\mathbb{Q}$, or some representation of $\mathbb{C}$). A differential indeterminate $\mathcal{D}$ is adjoined to form the *ring of differential polynomials* $\mathsf{F}(t)[\mathcal{D}; \delta]$, which consists of the polynomials in $\mathsf{F}(t)[\mathcal{D}]$ under the usual addition and a non-commutative multiplication defined such that $\mathcal{D}a = a\mathcal{D} + \delta(a)$, for any $a \in \mathsf{F}(t)$. Here $\delta : \mathsf{F}(t) \to \mathsf{F}(t)$ is a *pseudo-derivative*, a function such that for all $a, b \in \mathsf{F}(t)$ we have

$$\delta(a + b) = \delta(a) + \delta(b) \quad \text{and} \quad \delta(ab) = a\delta(b) + \delta(a)b.$$

The most common derivation in $\mathsf{F}(t)$ takes $\delta(a) = a'$ for any $a \in \mathsf{F}(t)$, the usual derivative of a, though other derivations (say $\delta(t) = t$) are certainly of interest.

A primary motivation in the definition of $\mathsf{F}(t)[D; \delta]$ is that there is a natural action on the space of infinitely differentiable functions in t, namely the differential polynomial

$$a_m \mathcal{D}^m + a_{m-1} \mathcal{D}^{m-1} + \cdots + a_1 \mathcal{D} + a_0 \in \mathsf{F}(t)[\mathcal{D}; \delta]$$

acts as the linear differential operator

$$a_m(t)\frac{d^m y(t)}{dt^m} + a_{m-1}(t)\frac{d^{m-1}y(t)}{dt^{m-1}} + \cdots + a_1(t)\frac{dy(t)}{dt} + a_0(t)y(t)$$

on a differentiable function $y(t)$. Solving and analyzing systems of such operators involves working with matrices over $\mathsf{F}(t)[\mathcal{D}; \delta]$, and invariants such as the differential analogues of the Smith, Popov and Hermite forms provide important structural information.

In commutative domains such as $\mathbb{Z}$ and $\mathsf{F}[x]$, it has been more common to compute the triangular Hermite and diagonal Smith form (as well as the lower degree Popov form, especially as an intermediate computation). Indeed, these forms are more canonical in the sense of being canonical in their class under multiplication by unimodular matrices. Polynomial-time algorithms for the Smith and Hermite forms over $\mathsf{F}[x]$ were developed by Kannan (1985), with important advances by Kaltofen et al. (1987), Villard (1995), Mulders and Storjohann (2003), and many others. One of the key features of this recent work in computing normal forms has been a careful analysis of the complexity in terms of matrix size, entry degree, and coefficient swell. Clearly identifying and analyzing the cost in terms of all these parameters has led to a dramatic drop in both theoretical and practical complexity.

Computing the classical Smith and Hermite forms of matrices over differential (and more general Ore) domains has received less attention though normal forms of differential polynomial matrices have applications in solving differential systems and control theory. Abramov and Bronstein (2001) analyzes the number of reduction steps necessary to compute a row-reduced form, while Beckermann et al. (2006) analyze the complexity of row reduction in terms of matrix size, degree and the sizes of the coefficients of some shifts of the input matrix. Beckermann et al. (2006) demonstrates tight bounds on the degree and coefficient sizes of the output, which we will employ here. For the Popov form, Cheng (2003) gives an algorithm for matrices of shift polynomials. Cheng's approach involves order bases computation in order to eliminate lower order terms of Ore polynomial matrices. A main contribution of Cheng (2003) is to give an algorithm computing the row rank and a row-reduced basis of the left nullspace of a matrix of Ore polynomials in a fraction-free way. This idea is extended in Davies et al. (2008) to compute Popov form of general Ore polynomial matrices. In Davies et al. (2008), they reduce the problem of computing Popov form to a nullspace computation. However, though Popov form is useful for rewriting high order terms with respect to low order terms, we want a different normal form more suited to solving system of linear diophantine equations. Since the Hermite form is upper triangular it meets this goal nicely, not to mention the fact that it is a "classical" canonical form. In a slightly different vein, Middeke (2008) has recently given an algorithm for the Smith (diagonal) form of a matrix of differential polynomials, which requires time polynomial in the matrix size and degree (but the coefficient size is not analyzed).

In this paper, we first discuss some basic operations with polynomials in $\mathsf{F}(t)[\mathcal{D}; \delta]$, which are typically written with respect to the differential variable $\mathcal{D}$ as

$$f = f_0 + f_1 \mathcal{D} + f_2 \mathcal{D}^2 + \cdots + f_d \mathcal{D}^d, \tag{1.1}$$

where $f_0, \ldots, f_d \in \mathsf{F}(t)$, with $f_d \neq 0$. We write $d = \deg_{\mathcal{D}} f$ to mean the degree in the differential variable, and generally refer to this as the *degree* of f. Since this is a non-commutative ring, it is important to set a standard notation in which the coefficients $f_0, \ldots, f_d \in \mathsf{F}(t)$ are written to the left of the differential variable $\mathcal{D}$. For $u, v \in \mathsf{F}[t]$ relatively prime, we can define $\deg_t(u/v) = \max\{\deg_t u, \deg_t v\}$. This is extended to $f \in \mathsf{F}(t)[\mathcal{D}; \delta]$ as in (1.1) by letting $\deg_t f = \max_i \{\deg_t f_i\}$. We think of $\deg_t$ as measuring coefficient size or height. Indeed, with a little extra work the bounds and algorithms in this paper are effective over $\mathbb{Q}(t)$ as well, where we also include the bit-length of rational coefficients, as well as the degree in t, in our analyses.

A matrix $U \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ is said to be *unimodular* if there exists a $V \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ such that $UV = I$, the $n \times n$ identity matrix. Note that we do not employ the typical determinantal definition of a unimodular matrix, as there is no easy notion of determinant for matrices over $\mathsf{F}(t)[\mathcal{D}; \delta]$ (indeed, working around this deficiency suffuses much of our work).

A matrix $H \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ is said to be in *Hermite form* if H is upper triangular, if every diagonal entry is monic, and every off-diagonal entry has degree less than the diagonal entry below it. As an example, the matrix

$$\begin{pmatrix} 1 + (t+2)\mathcal{D} + \mathcal{D}^2 & 2 + (2t+1)\mathcal{D} & 1 + (1+t)\mathcal{D} \\ 2t + t^2 + t\mathcal{D} & 2 + 2t + 2t^2 + \mathcal{D} & 4t + t^2 \\ 3 + t + (3+t)\mathcal{D} + \mathcal{D}^2 & 8 + 4t + (5+3t)\mathcal{D} + \mathcal{D}^2 & 7 + 8t + (2+4t)\mathcal{D} \end{pmatrix}$$

has Hermite form

$$\begin{pmatrix} 2 + t + \mathcal{D} & 1 + 2t & \frac{-2+t+2t^2}{2t} - \frac{1}{2t}\mathcal{D} \\ 0 & 2 + t + \mathcal{D} & 1 + \frac{7t}{2} + \frac{1}{2}\mathcal{D} \\ 0 & 0 & -\frac{2}{t} + \frac{-1+2t+t^2}{t}\mathcal{D} + \mathcal{D}^2 \end{pmatrix}.$$

Note that the Hermite form may have denominators in t. Also, while this example does not demonstrate it, it is common that the degrees in the Hermite form, in both t an $\mathcal{D}$, are substantially larger than in the input.

In this paper we will only concern ourselves with matrices in $\mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ of full row rank, that is, matrices whose rows are $\mathsf{F}(t)[\mathcal{D}; \delta]$-linear independent. For any matrix $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$, we show there exists a unimodular matrix U such that $UA = H$ is in Hermite form. This form is canonical in the sense that if two matrices $A, B \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ are such that $A = PB$ for unimodular $P \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ then the Hermite form of A equals the Hermite form of B.

The main contribution of this paper is an algorithm that, given a matrix $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ (of full row rank), computes H and U such that $UA = H$, which requires a polynomial number of F-operations in n, $\deg_{\mathcal{D}} A$, and $\deg_t A$. It will also require time polynomial in the coefficient bit-length when $\mathsf{F} = \mathbb{Q}$.

The remainder of the paper is organized as follows. In Section 2 we summarize some basic properties of differential polynomial rings and present and analyze algorithms for some necessary basic operations. In Section 3 we introduce a new approach to compute appropriate degree bounds on the coefficients of H and U. In Section 4 we present our algorithm for computing the Hermite form of a matrix of differential polynomials and analyze it completely.

2 Basic Structure and Operations in $\mathsf{F}[t][\mathcal{D}; \delta]$

In this section we discuss some of the basic structure of the ring $\mathsf{F}(t)[\mathcal{D}; \delta]$ and present and analyze simple algorithms to do some computations that will be necessary in the next section.

Some well-known properties of $\mathsf{F}(t)[\mathcal{D}; \delta]$ are worth recalling; see Bronstein and Petkovšek (1994) for an algorithmic presentation of this theory. Given $f, g \in \mathsf{F}(t)[\mathcal{D}; \delta]$, there is a degree function (in $\mathcal{D}$) which satisfies the usual properties: $\deg_{\mathcal{D}}(fg) = \deg_{\mathcal{D}} f + \deg_{\mathcal{D}} g$ and $\deg_{\mathcal{D}}(f + g) \leq \max\{\deg_{\mathcal{D}} f, \deg_{\mathcal{D}} g\}$. $\mathsf{F}(t)[\mathcal{D}; \delta]$ is also a left and right principal ideal ring, which implies the existence of a right (and left) division with remainder algorithm such that there exists unique $q, r \in \mathsf{F}(t)[\mathcal{D}; \delta]$ such that $f = qg + r$ where $\deg_{\mathcal{D}}(r) < \deg_{\mathcal{D}}(g)$. This allows for a right (and left) euclidean-like algorithm which shows the existence of a greatest common right divisor, $h = \gcd(f, g)$, a polynomial of minimal degree (in $\mathcal{D}$) such that $f = uh$ and $g = vh$ for $u, v \in \mathsf{F}(t)[\mathcal{D}; \delta]$. The GCRD is unique up to a left multiple in $\mathsf{F}(t) \backslash \{0\}$, and there exist co-factors $a, b \in \mathsf{F}(t)[\mathcal{D}; \delta]$ such that $af + bg = \gcd(f, g)$. There also exists a least common left multiple $\mathrm{lclm}(f, g)$. Analogously there exists a greatest common left divisor, $\gcd(f, g)$, and least common right multiple, $\mathrm{lcrm}(f, g)$, both of which are unique up to a right multiple in $\mathsf{F}(t)$.

Efficient algorithms for computing products of polynomials are developed in van der Hoeven (2002) and Bostan et al. (2008), while fast algorithms to compute the LCLM and GCRD, are developed in Li and Nemes (1997) and Li (1998). In this paper we will only need to compute very specific products of the form $\mathcal{D}^k f$ for some $k \in \mathbb{N}$. We will work with differential polynomials in $\mathsf{F}[t][\mathcal{D}; \delta]$, as opposed to $\mathsf{F}(t)[\mathcal{D}; \delta]$, and manage denominators separately. If $f \in \mathsf{F}[t][\mathcal{D}; \delta]$ is written as in (1.1), then $f_0, \ldots, f_d \in \mathsf{F}[t]$, and

$$\mathcal{D}f = \sum_{0 \leq i \leq d} f_i \mathcal{D}^{i+1} + \sum_{0 \leq i \leq d} f_i' \mathcal{D}^i \in \mathsf{F}[t][\mathcal{D}; \delta],$$

where $f_i' \in \mathsf{F}[t]$ is the usual derivative of $f_i \in \mathsf{F}[t]$. Assume $\deg_t f \leq e$. It is easily seen that $\deg_{\mathcal{D}}(\mathcal{D}f) = d + 1$, and $\deg_t(\mathcal{D}f) \leq e$. The cost of computing $\mathcal{D}f$ is $O(de)$ operations in F. Computing $\mathcal{D}^k f$, for $1 \leq k \leq m$ then requires $O(dem)$ operations in F.

If $\mathsf{F} = \mathbb{Q}$ we must account for the bit-length of the coefficients as well. Assuming our polynomials are in $\mathbb{Z}[t][\mathcal{D}; \delta]$ (which will be sufficient), and are written as above, we have $f_i = \sum_{0 \leq j \leq e} f_{ij} t^j$ for $f_{ij} \in \mathbb{Z}$. We write $\|f\|_\infty = \max |f_{ij}|$ to capture the coefficient size of f. It easily follows that $\|\mathcal{D}f\|_\infty \leq (e + 1)\|f\|_\infty$, and so $\|\mathcal{D}^m f\|_\infty \leq (e + 1)^m \|f\|_\infty$.

Lemma 2.1

(i) *Let $f \in \mathsf{F}[t][\mathcal{D};\delta]$ have $\deg_{\mathcal{D}} f = d$, $\deg_t f = e$, and let $m \in \mathbb{N}$. Then we can compute $\mathcal{D}^k f$, for $1 \leq k \leq m$, with $O(dem)$ operations in F.*

(ii) *Let $f \in \mathbb{Z}[t][\mathcal{D};\delta]$. Then $\|\mathcal{D}^m f\|_{\infty} \leq (e+1)^m \cdot \|f\|_{\infty}$, and we can compute $\mathcal{D}^i f$, for $1 \leq i \leq m$, with $O(dem \cdot (m \log e + \log \|f\|_{\infty})^2)$ bit operations.*

We make no claim that the above methods are the most efficient, and faster polynomial and matrix arithmetic will certainly improve the cost. However, the above analysis will be sufficient, and these costs will be dominated by others in the algorithms of later sections.

3   Existence and Degree Bounds on the Hermite Form

In this section we prove the existence and uniqueness of the Hermite form over $\mathsf{F}(t)[\mathcal{D};\delta]$, and prove some important properties about unimodular matrices and equivalence over this ring. The principal technical difficulty is that there is no natural determinant function with the properties found in commutative linear algebra. The determinant is one of the main tools used in the analysis of essentially all fast algorithms for computing the Hermite form H and transformation matrix U, and specifically two relevant techniques in established methods by Storjohann (1994) and Kaltofen et al. (1987). One approach might be to employ the non-commutative determinant of Dieudonné (1943), but this adds considerable complication. Instead, we find degree bounds via established bounds on the row-reduced form.

Definition 3.1 (Unimodular matrix). *Let $U \in \mathsf{F}(t)[\mathcal{D};\delta]^{n \times n}$ and suppose there exists a $V \in \mathsf{F}(t)[\mathcal{D};\delta]^{n \times n}$ such that $UV = I_n$, where I_n is the identity matrix over $\mathsf{F}(t)[\mathcal{D};\delta]^{n \times n}$. Then U is called a unimodular matrix over $\mathsf{F}(t)[\mathcal{D};\delta]$.*

This definition is in fact symmetric, in that V is also unimodular, as shown in the following lemma (the proof of which is left to the reader).

Lemma 3.1. *Let $U \in \mathsf{F}(t)[\mathcal{D};\delta]^{n \times n}$ be unimodular such that there exists a $V \in \mathsf{F}(t)[\mathcal{D};\delta]^{n \times n}$ with $UV = I_n$. Then $VU = I_n$ as well.*

Theorem 3.1. *Let $a, b \in \mathsf{F}(t)[\mathcal{D};\delta]$. There exists a unimodular matrix*

$$W = \begin{pmatrix} u & v \\ s & t \end{pmatrix} \in \mathsf{F}(t)[\mathcal{D};\delta]^{2 \times 2} \quad \text{such that} \quad W \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix},$$

where $g = \gcd(a, b)$ and $sa = -tb = \operatorname{lclm}(a, b)$.

Proof. Let $u, v \in \mathsf{F}(t)[\mathcal{D};\delta]$ be the multipliers from the euclidean algorithm such that $ua + vb = g$. Since $sa = -tb = \operatorname{lclm}(a, b)$, we know that $\gcd(s, t) = 1$ (otherwise the minimality of the degree of the lclm would be violated). It follows that there exist $c, d \in \mathsf{F}(t)[\mathcal{D};\delta]$ such that $sc + td = 1$. Now observe that

$$\begin{pmatrix} u & v \\ s & t \end{pmatrix} \begin{pmatrix} ag^{-1} & c \\ bg^{-1} & d \end{pmatrix} \begin{pmatrix} 1 & -uc - vd \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & uc + vd \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -uc - vd \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus
$$W^{-1} = \begin{pmatrix} ag^{-1} & ag^{-1}(-uc - vd) + c \\ bg^{-1} & bg^{-1}(-uc - vd) + d \end{pmatrix} = \begin{pmatrix} ag^{-1} & -a + c \\ bg^{-1} & -b + d \end{pmatrix},$$

so W is unimodular. □

Definition 3.2 (Hermite Normal Form). *Let $H \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ with full row rank. The matrix H is in Hermite form if H is upper triangular, if every diagonal entry of H is monic, and if every off-diagonal entry of H has degree (in $\mathcal{D}$) strictly lower than the degree of the diagonal entry below it.*

Theorem 3.2. *Let $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ have row rank n. Then there exists a matrix $H \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ with row rank n in Hermite form, and a unimodular matrix $U \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$, such that $UA = H$.*

Proof. We show this induction on n. The base case, $n = 1$, is trivial and we suppose that the theorem holds for $n - 1 \times n - 1$ matrices. Since A has row rank n, we can find a permutation of the rows of A such that every principal minor of A has full row rank. Since this permutation is a unimodular transformation of A, we assume this property about A. Thus, by the induction hypothesis, there exists a unimodular matrix $U_1 \in \mathsf{F}(t)[\mathcal{D}; \delta]^{(n-1) \times (n-1)}$ such that

$$\begin{pmatrix} & & & 0 \\ & U_1 & & 0 \\ & & & \vdots \\ & & & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \cdot A = \bar{H} = \begin{pmatrix} \bar{H}_{1,1} & \cdots & \cdots & & * & * \\ & \bar{H}_{2,2} & \cdots & & * & * \\ & & \ddots & & \vdots & \vdots \\ \mathbf{0} & & & \bar{H}_{n-1,n-1} & * \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n-1} & A_{n,n} \end{pmatrix} \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n},$$

where the $(n - 1)$st principal minor of $\bar{H}$ is in Hermite form. By Theorem 3.1, we know that there exists a unimodular matrix

$$W = \begin{pmatrix} u_i & v_i \\ s_i & -t_i \end{pmatrix} \in \mathsf{F}(t)[\mathcal{D}; \delta]^{2 \times 2} \quad \text{such that} \quad W \begin{pmatrix} \bar{H}_{ii} \\ A_{n,i} \end{pmatrix} = \begin{pmatrix} g_i \\ 0 \end{pmatrix} \in \mathsf{F}(t)[\mathcal{D}; \delta]^{2 \times 1}.$$

This allows us to reduce $A_{n,1}, \ldots, A_{n,n-1}$ to zero, and does not introduce any non-zero entries below the diagonal. Also, all off-diagonal entries can be reduced using unimodular operations modulo the diagonal entry, putting the matrix into Hermite form. □

Corollary 3.1. *Let $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ have full row rank. Suppose $UA = H$ for unimodular $U \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ and Hermite form $H \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$. Then both U and H are unique.*

Proof. Suppose H and G are both Hermite forms of A. Thus, there exist unimodular matrices U and V such that $UA = H$ and $VA = G$, and $G = WH$ where $W = VU^{-1}$ is unimodular. Since G and H are upper triangular matrices, we know W is as well. Moreover, since G and H have monic diagonal entries, the diagonal entries of W equal 1. We now prove W is the identity matrix. By

way of contradiction, first assume that W is not the identity, so there exists an entry W_{ij} which is the first nonzero off-diagonal entry on the ith row of W. Since $i < j$ and since $W_{ii} = 1$, $G_{ij} = H_{ij} + W_{ij}H_{jj}$. Because $W_{ij} \neq 0$, we see $\deg_{\mathcal{D}} G_{ij} \geq \deg_{\mathcal{D}} G_{jj}$, which contradicts the definition of the Hermite form. The uniqueness of U follows similarly.   □

Definition 3.3 (Row Degree). *A matrix $T \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ has row degree $\overrightarrow{u} \in (\mathbb{N} \cup \{-\infty\})^n$ if the ith row of T has degree u_i. We write $rowdeg\,\overrightarrow{u}$.*

Definition 3.4 (Leading Row Coefficient Matrix). *Let $T \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ have $rowdeg\,\overrightarrow{u}$. Set $N = \deg_{\mathcal{D}} T$ and $S = \mathrm{diag}(\mathcal{D}^{N-u_1}, \ldots, \mathcal{D}^{N-u_n})$. We write*

$$ST = L\mathcal{D}^N + \text{lower degree terms in } \mathcal{D},$$

where the matrix $L = LC_{row}(T) \in \mathsf{F}(t)^{n \times n}$ is called the leading row coefficient matrix of T.

Definition 3.5 (Row-reduced Form). *A matrix $T \in \mathsf{F}(t)[\mathcal{D}; \delta]^{m \times s}$ with rank r is in row-reduced form if rank $LC_{row}(T) = r$.*

Fact 3.1 (Beckermann et al. (2006) Theorem 2.2). *For any $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{m \times s}$ there exists a unimodular matrix $U \in \mathsf{F}(t)[\mathcal{D}; \delta]^{m \times m}$, with $T = UA$ having $r \leq \min\{m, s\}$ nonzero rows, $rowdeg\,T \leq rowdeg\,A$, and where the submatrix consisting of the r nonzero rows of T are row-reduced. Moreover, the unimodular multiplier satisfies the degree bound*

$$rowdeg\,U \leq \overrightarrow{v} + (|\overrightarrow{u}| - |\overrightarrow{v}| - \min_{j}\{u_j\})\overrightarrow{e},$$

where $\overrightarrow{u} := \max(\overrightarrow{0}, rowdeg\,A)$, $\overrightarrow{v} := \max(\overrightarrow{0}, rowdeg\,T)$, and $\overrightarrow{e}$ is the column vector with all entries equal to 1.

The proof of the following is left to the reader.

Corollary 3.2. *If $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ is a unimodular matrix then the row reduced form of A is an identity matrix.*

The following theorems provide degree bounds on H and U. We first compute a degree bound of the inverse of U by using the idea of backward substitution, and then use the result of Beckermann et al. (2006) to compute degree bound of U.

Theorem 3.3. *Let $A \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ be a matrix with $\deg_{\mathcal{D}} A_{ij} \leq d$ and full row rank. Suppose $UA = H$ for unimodular matrix $U \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ and $H \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ in Hermite form. Then there exist a unimodular matrix $V \in \mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$ such that $A = VH$ where $UV = I_n$ and $\deg_{\mathcal{D}} V_{ij} \leq d$.*

Proof. We prove by induction on n. The base case is $n = 1$. Since $H_{11} = \mathrm{gcrd}(A_{11}, \ldots, A_{n1})$, $\deg_{\mathcal{D}} H_{11} \leq d$ and so $\deg_{\mathcal{D}} V_{i1} \leq d$ for $1 \leq i \leq n$. Now,

we suppose that our claim is true for k where $1 < k < n$. Then we have to show that $\deg_{\mathcal{D}} V_{ik+1} \leq d$. We need to consider two cases:

Case 1: $\deg_{\mathcal{D}} V_{i,k+1} > \max(\deg_{\mathcal{D}} V_{i1}, \dots, \deg_{\mathcal{D}} V_{ik})$. Since

$$\deg_{\mathcal{D}} H_{k+1,k+1} \geq \max(\deg_{\mathcal{D}} H_{1,k+1}, \dots, \deg_{\mathcal{D}} H_{k,k+1}),$$
$$\deg_{\mathcal{D}} A_{i,k+1} = \deg_{\mathcal{D}}(V_{i,k+1} H_{k+1,k+1}),$$

where $A_{i,k+1} = V_{i1} H_{1,k+1} + \cdots + V_{i,k+1} H_{k+1,k+1}$. Thus, $\deg_{\mathcal{D}} V_{i,k+1} \leq d$.
Case 2: $\deg_{\mathcal{D}} V_{i,k+1} \leq \max(\deg_{\mathcal{D}} V_{i1}, \dots, \deg_{\mathcal{D}} V_{ik})$. Thus, by induction hypothesis, $\deg_{\mathcal{D}} V_{i,k+1} \leq d$. □

Corollary 3.3. *Let A, V, and U be those in Theorem 3.3. Then $\deg_{\mathcal{D}} U_{ij} \leq (n-1)d$.*

Proof. By Corollary 3.2, we know that the row reduced form of V is I_n. Moreover, since $I_n = UV$, we can compute the degree bound of U by using Fact 3.1. Clearly,

$$\overrightarrow{v} + (|\overrightarrow{u}| - |\overrightarrow{v}| - \min_j\{u_j\})\overrightarrow{e} \leq \overrightarrow{v} + (|\overrightarrow{u}| - \min_j\{u_j\})\overrightarrow{e},$$

where $\overrightarrow{u} := \max(\overrightarrow{0}, rowdegV)$ and $\overrightarrow{v} := \max(\overrightarrow{0}, rowdegI_n) = \overrightarrow{0}$. Since the degree of each row of V is bounded by d, $(|\overrightarrow{u}| - \min_j\{u_j\}) \leq (n-1)d$. Then, by Fact 3.1, $rowdegU \leq (n-1)d$. Therefore, $\deg_{\mathcal{D}} U_{ij} \leq (n-1)d$. □

Corollary 3.4. *Let H be same as that in Theorem 3.3. Then $\deg_{\mathcal{D}} H_{ij} \leq nd$.*

Proof. Since $\deg_{\mathcal{D}} U_{ij} \leq (n-1)d$ and $\deg_{\mathcal{D}} A_{ij} \leq d$, $\deg_{\mathcal{D}} H_{ij} \leq nd$. □

4 Computing Hermite Forms by Linear Systems over $\mathsf{F}(t)$

In this section we present our polynomial-time algorithm to compute the Hermite form of a matrix over $\mathsf{F}(t)[\mathcal{D}; \delta]$. We exhibit a variant of the linear system method developed in Kaltofen et al. (1987) and Storjohann (1994). The approach of these papers is to reduce the problem of computing the Hermite of matrices with (usual) polynomial entries in $\mathsf{F}[z]$ to the problem of solving a linear system equations over F. Analogously, we reduce the problem of computing the Hermite form over $\mathsf{F}[t][\mathcal{D}; \delta]$ to solving linear systems over $\mathsf{F}(t)$. The point is that the field $\mathsf{F}(t)$ over which we solve is the usual, commutative, field of rational functions.

For convenience, we assume that our matrix is over $\mathsf{F}[t][\mathcal{D}; \delta]$ instead of $\mathsf{F}(t)[\mathcal{D}; \delta]$, which can easily be achieved by clearing denominators with a "scalar" multiple from $\mathsf{F}[t]$. This is clearly a unimodular operation in the class of matrices over $\mathsf{F}(t)[\mathcal{D}; \delta]$.

We first consider formulating the computation of the Hermite form a matrix over $\mathsf{F}(t)[\mathcal{D}; \delta]$ as the solution of a "pseudo"-linear system over $\mathsf{F}(t)[\mathcal{D}; \delta]$ (i.e., a matrix equation over the non-commutative ring $\mathsf{F}(t)[\mathcal{D}; \delta]$).

Theorem 4.1. *Let $A \in \mathsf{F}[t][\mathcal{D}; \delta]^{n \times n}$ have full row rank, with $\deg_{\mathcal{D}} A_{i,j} \leq d$, and $(d_1, \dots, d_n) \in \mathbb{N}^n$ be given. Consider the system of equations $PA = G$, for $n \times n$ matrices for $P, G \in \mathsf{F}(t)[\mathcal{D}; \delta]$ restricted as follows:*

- *The degree (in $\mathcal{D}$) of each entry of P is bounded by $(n-1)d + \max_{1 \le i \le n} d_i$.*
- *The matrix G is upper triangular, where every diagonal entry is monic and the degree of each off-diagonal entry is less than the degree of the diagonal entry below it.*
- *The degree of the ith diagonal entry of G is d_i.*

Let H be the Hermite form of A and $(h_1, \ldots, h_n) \in \mathbb{N}^n$ be the degrees of the diagonal entries of H. Then the following are true:

(a) *There exists at least one pair P, G as above with $PA = G$ if and only if $d_i \ge h_i$ for $1 \le i \le n$.*
(b) *If $d_i = h_i$ for $1 \le i \le n$ then G is the Hermite form of A and P is a unimodular matrix.*

Proof. The proof is similar to that of Kaltofen et al. (1987), Lemma 2.1. Given a degree vector $(d_1, \ldots, d_n)$, we view $PA = G$ as a system of equations in the unknown entries of P and G. Since H is the Hermite form of A, there exist a unimodular matrix U such that $UA = H$. Thus $PU^{-1}H = G$ and the matrix PU^{-1} must be upper triangular since the matrices H and G are upper triangular. Moreover, since the matrix PU^{-1} is in $\mathsf{F}(t)[\mathcal{D}; \delta]^{n \times n}$, and $G_{ii} = (PU^{-1})_{ii} \cdot H_{ii}$ for $1 \le i \le n$, we know $d_i \ge h_i$ for $1 \le i \le n$. For the other direction, we suppose $d_i \ge h_i$ for $1 \le i \le n$. Let $D = \mathrm{diag}(\mathcal{D}^{d_1 - h_1}, \ldots, \mathcal{D}^{d_n - h_n})$. Then since $(DU)A = (DH)$, we can set $P = DU$ and $G = DH$ as a solution to $PA = G$, and the ith diagonal of G has degree d_i by construction. By Corollary 3.3, we know $\deg_{\mathcal{D}} U_{i,j} \le (n-1)d$ and so $\deg_{\mathcal{D}} P_{i,j} \le (n-1)d + \max_{1 \le i \le n} d_i$.

To prove (b), suppose $d_i = h_i$ for $1 \le i \le n$ and that, contrarily, G is *not* the Hermite form of A. Since PU^{-1} is an upper triangular matrix with ones on the diagonal, PU^{-1} is a unimodular matrix. Thus P is a unimodular matrix and, by Corollary 3.1, G *is* the (unique) Hermite form of A, a contradiction.   □

Lemma 4.1. *Let A, P, $(d_1, \ldots, d_n)$, and G be as in Theorem 4.1, and let $\beta := (n-1)d + \max_{1 \le i \le n} d_i$. Also, assume that $\deg_t A_{ij} \le e$ for $1 \le i, j \le n$. Then we can express the system $PA = G$ as a linear system over $\mathsf{F}(t)$ as $\widehat{P}\widehat{A} = \widehat{G}$ where*

$$\widehat{P} \in \mathsf{F}(t)^{n \times n(\beta+1)}, \quad \widehat{A} \in \mathsf{F}[t]^{n(\beta+1) \times n(\beta+d+1)}, \quad \widehat{G} \in F(t)^{n \times n(\beta+d+1)}.$$

Assuming the entries $\widehat{A}$ are known while the entries of $\widehat{P}$ and $\widehat{G}$ are indeterminates, the system of equations from $\widehat{P}\widehat{A} = \widehat{G}$ for the entries of $\widehat{P}$ and $\widehat{G}$ is linear over $\mathsf{F}(t)$ in its unknowns, and the number of equations and unknowns is $O(n^3 d)$. The entries in $\widehat{A}$ are in $\mathsf{F}[t]$ and have degree at most e.

Proof. Since $\deg_{\mathcal{D}} P_{i,j} \le \beta$, each entry of P has at most $(\beta+1)$ coefficients in $\mathsf{F}(t)$ and can be written as $P_{ij} = \sum_{0 \le k \le \beta} P_{ijk}\mathcal{D}^k$. We let $\widehat{P} \in \mathsf{F}(t)^{n \times n(\beta+1)}$ be the matrix formed from P with P_{ij} replaced by the row vector $(P_{ij0}, \ldots, P_{ij\beta}) \in \mathsf{F}(t)$.

Since $\deg_{\mathcal{D}} P \le \beta$, when forming PA, the entries in A are multiplied by $\mathcal{D}^\ell$ for $0 \le \ell \le \beta$, resulting in polynomials of degree in $\mathcal{D}$ of degree at most $\mu = \beta + d$.

Thus, we construct $\widehat{A}$ as the matrix formed from A with A_{ij} replaced by the $(\beta + 1) \times (\mu + 1)$ matrix whose ℓth row is

$$(A_{ij0}^{[\ell]}, A_{ij1}^{[\ell]}, \ldots, A_{ij\mu}^{[\ell]}) \quad \text{such that} \quad \mathcal{D}^\ell A_{ij} = A_{ij0}^{[\ell]} + A_{ij1}^{[\ell]} \mathcal{D} + \cdots + A_{ij\mu}^{[\ell]} \mathcal{D}^\mu.$$

Note that by Lemma 2.1 we can compute $\mathcal{D}^\ell A_{i,j}$ quickly.

Finally, we construct the matrix $\widehat{G}$. Each entry of G has degree in $\mathcal{D}$ of degree at most $nd \le n(\beta + d + 1)$. Thus, initially $\widehat{G}$ is the matrix formed by G with G_{ij} replaced by

$$(G_{ij0}, \ldots, G_{ij\mu}) \quad \text{where} \quad G_{ij} = G_{ij0} + G_{ij1}\mathcal{D} + \cdots + G_{ij\mu}\mathcal{D}^\mu.$$

However, because of the structure of the system we can fix values of many of the entries of $\widehat{G}$ as follows. First, since every diagonal entry of the Hermite form is monic, we know the corresponding entry in $\widehat{G}$ is 1. Also, by Corollary 3.4, the degree in $\mathcal{D}$ of every diagonal entry of H is bounded by nd, and every off-diagonal has degree in $\mathcal{D}$ less than that of the diagonal below it (and hence less than nd), and we can set all coefficients of larger powers of $\mathcal{D}$ to 0 in $\widehat{G}$.

The resulting system $\widehat{P}\widehat{A} = \widehat{G}$, restricted as above according to Theorem 4.1, has $O(n^3d)$ linear equations in $O(n^3d)$ unknowns. Since the coefficients in $\widehat{A}$ are all of the form $\mathcal{D}^\ell A_{ij}$, and since this does not affect their degree in t, the degree in t of entries of $\widehat{A}$ is the same as that of A, namely e. $\qquad\square$

With more work, we believe the dimension of the system can be reduced to $O(n^2d) \times O(n^2d)$ if we apply the techniques presented in Storjohann (1994) Section 4.3, wherein the unknown coefficients of $\widehat{G}$ are removed from the system. See also Labhalla et al. (1996).

So far, we have shown how to convert the differential system over $\mathsf{F}(t)[\mathcal{D}; \delta]$ into a linear system over $\mathsf{F}(t)$. Also, we note, by Theorem 4.1, that the correct degree of the ith diagonal entry in the Hermite form of A can be found by seeking the smallest non-negative integer k such that $PA = G$ is consistent when $\deg_\mathcal{D} G_{j,j} = nd$ for $j = 1, \ldots, i - 1, i + 1, \ldots, n$ and $k \le \deg_\mathcal{D} G_{i,i}$. Using binary search, we can find the correct degrees of all diagonal entries by solving at most $O(n \log(nd))$ systems. We then find the correct degrees of the diagonal entries in the Hermite form of A, solving the system $PA = G$ with the correct diagonal degrees gives the matrices U and H such that $UA = H$ where H is the Hermite form of A.

Theorem 4.2. *Let* $A \in \mathsf{F}[t][\mathcal{D}; \delta]^{n \times n}$ *with* $\deg_\mathcal{D} A_{ij} \le d$ *and* $\deg_t A_{ij} \le e$ *for* $1 \le i, j \le n$. *Then we can compute the Hermite form* $H \in \mathsf{F}(t)[\mathcal{D}; \delta]$ *of* A, *and a unimodular* $U \in \mathsf{F}[t][\mathcal{D}; \delta]$ *such that* $UA = H$, *with* $O((n^{10}d^3 + n^7d^2e) \log(nd))$ *operations in* F

Proof. Lemma 4.1 and the following discussion, above shows that computing U and H is reduced to solving $O(n \log(nd))$ systems of linear equations over $\mathsf{F}(t)$, each of which is $m \times m$ for $m = O(n^3d)$ and in which the entries have degree e. Using standard linear algebra this can be solved with $O(m^4e)$ operations in F, since any solution has degree at most me (see von zur Gathen and Gerhard

(2003)). A somewhat better strategy is to use the t-adic lifting approach of Dixon (1982), which would require $O(m^3 + m^2 e)$ operations in F for each system, giving a total cost of $O((n^{10}d^3 + n^7 d^2 e)\log(nd))$ operations in F.   □

As noted above, it is expected that we can bring this cost down through a smaller system similar to that of Storjohann (1994), to a cost of $O((n^7 d^2 + n^5 d^2 e)\log(nd))$. Nonetheless, the algorithm as it is stated achieves a guaranteed polynomial-time solution.

   It is often the case that we are considering differential systems over $\mathbb{Q}(t)[\mathcal{D}; \delta]$, where we must contend with growth in coefficients in $\mathcal{D}$, t *and* in the size of the rational coefficients. However, once again we may employ the fact that the Hermite form and unimodular transformation matrix are solutions of a linear system over $\mathbb{Q}[t]$. For convenience, we can assume in fact that our input is in $\mathbb{Z}[t][\mathcal{D}; \delta]^{n \times n}$ (since the rational matrix to eliminate denominators is unimodular in $\mathbb{Q}(t)[\mathcal{D}; \delta]$). There is some amount of extra coefficient growth when going from A to $\widehat{A}$; namely we take up to nd derivatives, introducing a multiplicative constant of size around $\min((nd)!, e!)$. In terms of the bit-length of the coefficients, this incurs a multiplicative blow-up of only $O(\ell \log(\ell))$ where $\ell = \min(nd, e)$. It follows that we can find the Hermite form of $A \in \mathbb{Q}(t)[\mathcal{D}; \delta]^{n \times n}$ in time polynomial in n, $\deg_t A_{ij}$, $\deg_{\mathcal{D}} A_{ij}$, and $\log \|A_{ij}\|$, the maximum coefficient length in an entry, for $1 \leq i, j \leq n$. A modular algorithm, for example along the lines of Li and Nemes (1997), would improve performance considerably, as might p-adic solvers and a more careful construction of the linear system.

5   Conclusions and Future Work

We have shown that the problem of computing the Hermite form of a matrix over $\mathsf{F}(t)[\mathcal{D}; \delta]$ can be accomplished in polynomial time. Moreover, our algorithm will also control growth in coefficient bit-length when $\mathsf{F} = \mathbb{Q}$. We have also shown that the degree bounds on Hermite forms in the differential ring are very similar to the regular polynomial case. From a practical point of view our method is still expensive. Our next work will be to investigate more efficient algorithms. We have suggested ways to compress the system of equations and to employ structured matrix techniques. Also, the use of randomization has been shown to be highly beneficial over $\mathsf{F}[t]$, and should be investigated in this domain. Finally, our approach should be applicable to difference polynomials and more general Ore polynomial rings.

References

Abramov, S., Bronstein, M.: On solutions of linear functional systems. In: Proc. ACM International Symposium on Symbolic and Algebraic Computation, pp. 1–7 (2001)

Beckermann, B., Cheng, H., Labahn, G.: Fraction-free row reduction of matrices of ore polynomials. Journal of Symbolic Computation 41(1), 513–543 (2006)

Bostan, A., Chyzak, F., Le Roux, N.: Products of ordinary differential operators by evaluation and interpolation. In: Proc. International Symposium on Symbolic and Algebraic Computation, pp. 23–30 (2008)

Bronstein, M., Petkovšek, M.: On Ore rings, linear operators and factorisation. Programmirovanie 20, 27–45 (1994)

Cheng, H.: Algorithms for Normal Forms for Matrices of Polynomials and Ore Polynomials. PhD thesis, University of Waterloo (2003),
http://www.cs.uleth.ca/~cheng/publications.html

Davies, P., Cheng, H., Labahn, G.: Computing Popov form of general Ore polynomial matrices. In: Milestones in Computer Algebra, pp. 149–156 (2008)

Dickson, L.E.: Algebras and their arithmetics. G.E. Stechert, New York (1923)

Jean Dieudonné, M.: Les déterminants sur un corps non commutatif. Bulletin de la Société Mathématique de France 71, 27–45 (1943)

Dixon, J.D.: Exact solution of linear equations using p-adic expansions. Numer. Math. 40, 137–141 (1982)

von zur Gathen, J., Gerhard, J.: Modern Computer Algebra. Cambridge University Press, Cambridge (2003)

Hermite, C.: Sur les fonctions de sept lettres. C.R. Acad. Sci. Paris 57, 750–757 (1863); OE uvres. Gauthier-Villars, Paris 2, 280–288 (1908)

van der Hoeven, J.: FFT-like multiplication of linear differential operators. Journal of Symbolic Computation 33(1), 123–127 (2002)

Jacobson, N.: The Theory of Rings. American Math. Soc., New York (1943)

Kaltofen, E., Krishnamoorthy, M.S., Saunders, B.D.: Fast parallel computation of Hermite and Smith forms of polynomial matrices. SIAM J. Algebraic and Discrete Methods 8, 683–690 (1987)

Kannan, R.: Polynomial-time algorithms for solving systems of linear equations over polynomials. Theoretical Computer Science 39, 69–88 (1985)

Labhalla, S., Lombardi, H., Marlin, R.: Algorithmes de calcul de la réduction de Hermite d'une matrice à coefficients polynomiaux. Theoretical Computer Science 161(1–2), 69–92 (1996)

Li, Z.: A subresultant theory for Ore polynomials with applications. In: Proc. International Symposium on Symbolic and Algebraic Computation, pp. 132–139 (1998)

Li, Z., Nemes, I.: A modular algorithm for computing greatest common right divisors of ore polynomials. In: ISSAC 1997: Proceedings of the 1997 international symposium on Symbolic and algebraic computation. ACM, New York (1997)

Middeke, J.: A polynomial-time algorithm for the jacobson form for matrices of differential operators. Technical Report 08-13, Research Institute for Symbolic Computation (RISC), Linz, Austria (2008)

Mulders, T., Storjohann, A.: On lattice reduction for polynomial matrices. Journal of Symbolic Computation 35(4), 377–401 (2003)

Popov, V.: Invariant description of linear, time-invariant controllable systems. SIAM J. Control 10, 252–264 (1972)

Smith, H.J.S.: On systems of linear indeterminate equations and congruences. Philos. Trans. Royal Soc. London 151, 293–326 (1861)

Storjohann, A.: Computation of Hermite and Smith normal forms of matrices. Master's thesis, University of Waterloo (1994)

Villard, G.: Generalized subresultants for computing the smith normal form of polynomial matrices. Journal of Symbolic Computation 20, 269–286 (1995)

Wedderburn, J.H.M.: Non-commutative domains of integrity. Journal für die reine und angewandte Mathematik 167, 129–141 (1932)