Quasideterminants, Degree Bounds and "Fast" Algorithms for Matrices of Ore Polynomials

Mark Giesbrecht

with Albert Heinle (Sortable)
Myung Sub Kim (VMWare)

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

November 21, 2017

Ore Polynomials – Definition and Notation

Definition (Ore Polynomials)

Let F be a skew field

- $\sigma : \mathsf{F} \to \mathsf{F}$ an automorphism
- $\delta : \mathsf{F} \to \mathsf{F}$ a σ-derivation: For all $a, b \in \mathsf{F}$
  $\delta(a+b) = \delta(a) + \delta(b)$ and $\delta(ab) = \sigma(a)\delta(b) + \delta(a)b$

Define $\mathsf{F}[X;\sigma,\delta]$ as a ring of polynomials in $\mathsf{F}[X]$

- Usual polynomial addition $(+)$
- Multiplication: $Xa = \sigma(a)X + \delta(a)$ for any $a \in \mathsf{F}$

Prototypical examples: $\mathsf{F} = \mathsf{K}(t)$ for a field K

- $\sigma(t) = t + 1$, $\delta(t) = 0$
  ➡ $Xt = (t+1)X$ the *shift polynomials*
- $\sigma(t) = t$, $\delta(t) = 1$
  ➡ $Xf(t) = f(t)X + \frac{d}{dt}f(t)$ the *differential polynomials*

Why Ore polynomials?

- Defined by Ore (1933,1934) as a concrete unification of linear differential, and difference equations.
- Left (and right) principal ideal/euclidean domain
- Well-behaved degree function $\deg_X$
- Applications to solving systems of linear differential, difference equations, finite fields
- “Base case” for multivariate non-commutative polynomial rings

Why Ore polynomials?

- Defined by Ore (1933,1934) as a concrete unification of linear differential, and difference equations.
- Left (and right) principal ideal/euclidean domain
- Well-behaved degree function $\deg_X$
- Applications to solving systems of linear differential, difference equations, finite fields
- “Base case” for multivariate non-commutative polynomial rings

Why Matrices of Ore polynomials?

- Systems of linear differential and difference operators
- Determining invariants of these systems

Canonical matrix forms over $\mathsf{F}[X;\sigma,\delta]$

The Euclidean Domain structure of $\mathsf{F}[X;\sigma,\delta]$ gives a rich structure to the matrices over $\mathsf{F}[X;\sigma,\delta]$.

Definition (Hermite canonical form)

$H \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ is in *Hermite form* if

- H is upper triangular
- diagonal elements are monic (i.e., leading term 1)
- $\deg H_{ij} < \deg H_{jj}$ for $1 \leqslant i < j \leqslant n$, (i.e., each diagonal entry of higher degree than entries above it).

Theorem

- *For every $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ there exists a unimodular $U \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ such that $H = UA$ is in Hermite form.*
- *The Hermite form is unique.*

Hermite form example

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Hermite form:

$$\text{Let } U = \begin{bmatrix} \frac{1-t}{2t} & \frac{1}{t}+\frac{1}{2t}X & -\frac{1}{2t} \\ \frac{t}{2}-\frac{1}{2}X & -\frac{1}{2}X & \frac{1}{2} \\ \frac{1+2t^2}{t}+(t-1)X & \frac{2}{t}+\frac{1-2t}{t}X-X^2 & -\frac{1}{t}+X \end{bmatrix}$$

$$\text{Then } UA = H = \begin{bmatrix} 2+t+X & 1+2t & \frac{-2+t+2t^2}{2t}-\frac{1}{2t}X \\ 0 & 2+t+X & 1+\frac{7t}{2}+\frac{1}{2}X \\ 0 & 0 & -\frac{2}{t}+\frac{-1+2t+t^2}{t}X+X^2 \end{bmatrix}$$

Growth in all directions:

Want efficiency in terms of n, $\deg_X A$, $\deg_t(A)$ and $\log|A_{ij}|$

Canonical matrix forms over $\mathsf{F}[X;\sigma,\delta]$

Definition: Jacobson form

$S \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ in *Jacobson* form iff

- $S = \mathrm{diag}(s_1,\ldots,s_n) \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$
- $s_i \in \mathsf{F}[X;\sigma,\delta]$ is a left and right — *total* — divisor of s_{i+1}

Canonical matrix forms over $\mathsf{F}[X;\sigma,\delta]$

Definition: Jacobson form

$S \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ in *Jacobson* form iff

- $S = \mathrm{diag}(s_1,\ldots,s_n) \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$
- $s_i \in \mathsf{F}[X;\sigma,\delta]$ is a left and right — *total* — divisor of s_{i+1}

Theorem

For every $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ there exist unimodular $U, V \in \mathsf{F}[X;\sigma,\delta]$ such that UAV is in Jacobson form.

- Unimodular means invertible over $\mathsf{F}[X;\sigma,\delta]$
- Diagonal entries of Jacobson form unique up to *similarity*: $f, g \in \mathsf{F}[X;\sigma,\delta]$ are *similar* if there exists $u \in \mathsf{F}[X;\sigma,\delta]$ with $\mathrm{gcrd}(u,f) = 1$ and $g = \mathrm{lclm}(u,f)\cdot u^{-1}$

Canonical matrix forms over $\mathsf{F}[X;\sigma,\delta]$

Definition: Jacobson form

$S \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ in *Jacobson* form iff

- $S = \operatorname{diag}(s_1,\ldots,s_n) \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$
- $s_i \in \mathsf{F}[X;\sigma,\delta]$ is a left and right — *total* — divisor of s_{i+1}

Stronger characterization for differential polynomials

Theorem

Let $A \in \mathbb{Q}(t)[X;\delta]$ have full row rank, where $Xt = tX + 1$ (differential polynomials). Then A has Jacobson form

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \psi \end{pmatrix} \in \mathbb{Q}(t)[X;\delta]^{n\times n},$$

for some $\psi \in \mathbb{Q}(t)[X;\delta]$

Canonical matrix forms over $\mathsf{F}[X;\sigma,\delta]$

Definition: Jacobson form

$S \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ in *Jacobson* form iff

- $S = \operatorname{diag}(s_1,\ldots,s_n) \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$
- $s_i \in \mathsf{F}[X;\sigma,\delta]$ is a left and right — *total* — divisor of s_{i+1}

Stronger characterization for shift polynomials:

Theorem

Let $A \in \mathbb{Q}(t)[X;\sigma]$ have full row rank, where $Xt = (t+1)X$ (shift polynomials). Then A has Jacobson form

$$\begin{pmatrix} X^{j_1} & & & \\ & \ddots & & \\ & & X^{j_{n-1}} & \\ & & & \varphi(X)X^{j_n} \end{pmatrix} \in \mathbb{Q}(t)[X;\sigma]^{n\times n} \qquad j_1 \leqslant j_2 \leqslant \cdots \leqslant j_n$$

for some $\varphi \in \mathbb{Q}(t)[X;\sigma]$ such that $\operatorname{gcrd}(\varphi, X) = 1$.

An Example: Jacobson (differential)

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Jacobson form:

There exist unimodular matrices $U, V \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with

$$UAV = J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \begin{array}{l} \left(\frac{-2(t+2)^2}{t}\right) + \left(\frac{11t^2+6t^3+t^4-12}{t}\right)X + \\ \quad + \left(\frac{12t^2+3t^3+10t-6}{t}\right)X^2 + \left(\frac{3t^2+6t-1}{t}\right)X^3 + X^4 \end{array} \end{bmatrix}$$

Growth in all directions:

Want efficiency in terms of n, $\deg_X(A)$, $\deg_t(A)$ and $\log|A_{ij}|$

Commutative analogues

Jacobson and Hermite forms have analogues over $\mathbb{Z}$ and $\mathbb{Q}[x]$.
Hermite, and especially Smith form are common in number-theoretic and polynomial computations.

Canonical forms over F$[x]$

$$A = \begin{pmatrix} -2+2x & 2x+2 & 4x-6 \\ 2x^2-2 & -2x^2+4x-2 & 4x^2-14x+10 \\ 4x^2-10x+6 & -2x^2-12+2x^3 & 19x^2-65x+52 \end{pmatrix}$$

$$\Rightarrow UA = H = \begin{pmatrix} x-1 & x+1 & 2\,x-3 \\ 0 & x^2+1 & 3\,x-4 \\ 0 & 0 & x^2-3\,x+2 \end{pmatrix}$$

Commutative analogues

Jacobson and Hermite forms have analogues over $\mathbb{Z}$ and $\mathbb{Q}[x]$.
Hermite, and especially Smith form are common in number-theoretic and polynomial computations.

Canonical forms over F$[x]$

$$A = \begin{pmatrix} -2+2x & 2x+2 & 4x-6 \\ 2x^2-2 & -2x^2+4x-2 & 4x^2-14x+10 \\ 4x^2-10x+6 & -2x^2-12+2x^3 & 19x^2-65x+52 \end{pmatrix}$$

$$\Rightarrow UAV = S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & x^4-3\,x^3+3\,x^2-3\,x+2 \end{pmatrix}$$

Hermite/Smith over $\mathbb{Z}$ & $\mathsf{F}[x]$: a complexity success story

Let $A \in \mathsf{F}[x]^{n\times n}$, where $\deg_x A \leqslant d$, $\text{sizeof}(A_{ij}) = |A_{ij}| \leqslant \beta$.
Find $U \in \mathsf{F}[x]^{n\times n}$, $H \in \mathsf{F}[x]$ in Hermite form such that $UA = H$.

- Hermite (1851): exponential time
- Kannan (1985): $(nd)^{O(1)}$
- Kaltofen, Krishnamurthy, & Saunders (1987): $(nd \cdot \log\beta)^{O(1)}$
- Storjohann & Labahn (1995): $O(n^5 d \log(\beta)(d + \log\beta))$
- Storjohann & Mulders (2003): $O(n^3 d \log(\beta)(d + \log\beta))$

Now also the fastest algorithms in practice

Tools

- Randomization
- Determinantal bounds
- "linearization"
- Restricted Gröbner bases ➡ Popov form

Canonical forms over $\mathsf{F}[X;\sigma,\delta]$: State of the Art

Let $B \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$. Think of B as a matrix polynomial

$$B = B_0 + B_1 X + B_2 X^2 + \cdots + B_d X^d, \quad B_i \in \mathsf{F}^{n\times n}.$$

B is in row-reduced form if the $\operatorname{rank} B_d = \operatorname{rank} B$.
For $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ there exists unimodular $U \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ such that UA is row reduced.

- Row reduction reveals rank, useful for reducing order of system
- Abramov & Bronstein (2001) compute a rank-revealing transformation and analyze the number of reduction steps
- Beckermann, Cheng & Labahn (2006) for row reduced form with tight bounds on various row degrees:
  Given $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$, with sizeof$(A_{ij}) \leqslant \beta$ their algorithm requires time polynomial in $(n + \deg A + \beta)^{O(1)}$

Linear Algebra over $\mathsf{F}[X;\sigma,\delta]$: State of the Art

Popov form

The Popov (1969) form is another canonical form useful because it maintains low degree (but is not triangular)

- Davies, Cheng, Labahn (2008) compute Popov form of general Ore polynomial matrices (prove some degree bounds)

Linear Algebra over $\mathsf{F}[X;\sigma,\delta]$: State of the Art

Popov form

The Popov (1969) form is another canonical form useful because it maintains low degree (but is not triangular)

- Davies, Cheng, Labahn (2008) compute Popov form of general Ore polynomial matrices (prove some degree bounds)

Jacobson and Hermite form Computation

- Blinkov, Cid, Gerdt, Plesken, Robertz (2003): implementation of Jacobson form in Janet.
- Culianez & Quadrat (2005): Jacobson and Hermite
- Levandovskyy & Schindelar (2010, 2011): Jacobson via GB

Linear Algebra over $\mathsf{F}[X;\sigma,\delta]$: State of the Art

Popov form

The Popov (1969) form is another canonical form useful because it maintains low degree (but is not triangular)

- Davies, Cheng, Labahn (2008) compute Popov form of general Ore polynomial matrices (prove some degree bounds)

Jacobson and Hermite form Computation

Middeke (2008,2011): Jacobson form of a $A \in \mathsf{F}[\mathcal{D};\delta]^{n\times n}$

- Different method using cyclic vectors.
- Polynomial time in n and $d = \deg A$: $O(n^8 d^3)$ operations in F
- Conversion of Popov to Hermite using FGLM

Linear Algebra over $\mathsf{F}[X;\sigma,\delta]$: State of the Art

Popov form

The Popov (1969) form is another canonical form useful because it maintains low degree (but is not triangular)

- Davies, Cheng, Labahn (2008) compute Popov form of general Ore polynomial matrices (prove some degree bounds)

Fast Popov Form Computation

Khochtali, Rosenkilde, Storjohann (ISSAC'17)

- Compute Popov form of $A \in \mathsf{K}[t][X;\sigma,\delta]^{n\times n}$
- Cost $O(n^4 d^3 e)$ where $d = \deg_X A$ and $e = \deg_t A$

A Computational View of Ore Polynomials

Ground field F

Let F be a (*not necessarily commutative*) field.
Assume F has a *size* function sizeof : $F \to \mathbb{N}$ such that for $\alpha, \beta \in F$

- $\text{sizeof}(\alpha\beta) \in \tilde{O}(\text{sizeof}(\alpha) + \text{sizeof}(\beta))$
- $\text{sizeof}(\alpha + \beta) \in \tilde{O}(\text{sizeof}(\alpha) + \text{sizeof}(\beta))$
- $\text{sizeof}(\alpha^{-1}) = \text{sizeof}(\alpha)$
- $\text{sizeof}(\sigma(\alpha)) \in \tilde{O}(\text{sizeof}(\alpha))$, $\quad \text{sizeof}(\delta(\alpha)) \in \tilde{O}(\text{sizeof}(\alpha))$

More stringent or relaxed specs will yield analogous results.

Efficient linear algebra in F

Assumption: Given $B \in F^{m\times n}$, $b \in F^{n\times 1}$

- Solve $Bv = b$ for $b \in F^{n\times 1}$ (or show no such solution exists)
- Determine $\mathrm{rank}\, B$

with $\tilde{O}(n^2 m\beta)$ operations in F, where $\beta = \max_{ij} \text{sizeof}(B_{ij})$.

Degree Bounds for Hermite forms

Determinants: A Missing Tool

A primary tool in the commutative case for bounding the output size is the *determinant*. Not available for skew fields (?)

Dieudonné determinant

Let E be any skew field

For $A \in \mathsf{E}^{n \times n}$, find Bruhat factorization of $A = PLDU$:

- $P \in \mathsf{E}^{n \times n}$ a permutation matrix
- L, $U \in \mathsf{E}^{n \times n}$ lower/upper triangular, 1 on diagonal
- $D = \mathrm{diag}(u_1, \ldots, u_n) \in \mathsf{E}^{n \times n}$

Define $\delta\varepsilon\tau(A) \equiv u_1 \cdots u_n \bmod [\mathsf{E}^*, \mathsf{E}^*]$

Dieudonné determinant over $\mathsf{F}[X;\sigma,\delta]$

For $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$, find Bruhat factorization of $A = PLDU$:

- $P \in \mathsf{F}^{n\times n}$ a permutation matrix
- $L, U \in \mathsf{F}(X;\sigma,\delta)^{n\times n}$ lower/upper triangular, 1 on diagonal
- $D = \mathrm{diag}(u_1,\ldots,u_n) \in \mathsf{F}(X;\sigma,\delta)^{n\times n}$

Define $\delta\varepsilon\tau(A) \equiv u_1 \cdots u_n \bmod [\mathsf{F}[X;\sigma,\delta]^*, \mathsf{F}[X;\sigma,\delta]^*]$

Nice properties of the Dieudonné determinant

- Multiplicative: $\delta\varepsilon\tau(AB) = \delta\varepsilon\tau(A)\cdot\delta\varepsilon\tau(B)$
- $\deg\delta\varepsilon\tau(AB) = \deg\delta\varepsilon\tau(A) + \deg\delta\varepsilon\tau(B)$ (Taelman, 2006)

Deficiencies of the Dieudonné determinant

- No Cramer's rule, Leibniz formula, or ability to bound degrees.

Quasideterminants

Gelfand & Retakh (1991) define **quasideterminant(s)**.

> *We believe that the notion of quasideterminants should be one of main organizing tools in noncommutative algebra giving them the same role determinants play in commutative algebra.*

Let $A \in \mathsf{E}^{n\times n}$ over a skew field E, and $B = A^{-1}$

Define the (p, q) quasideterminant of A:

$$\det_{pq} A = \frac{1}{(A^{-1})_{qp}}$$

Recursive expansion:

$$\det_{pq}(A) = A_{pq} - \sum_{i\neq p, j\neq q} A_{pi}(\det_{ji}(A^{(pq)}))^{-1} A_{jq}$$

where $A^{(pq)}$ is A with row p and column q removed.

- Some entries may be undefined!

Degree bounds and quasideterminants over $\mathsf{F}[X;\sigma,\delta]$

Need to extend degree function naturally to quotient skew field $\mathsf{F}(X;\sigma,\delta)$:

- Any $h \in \mathsf{F}(X;\sigma,\delta)$ can be written as $u \cdot v^{-1}$ for $u, v \in \mathsf{F}[X;\sigma,\delta]$ (non-unique)
- Define: $\deg h := \deg u - \deg v$

For any $h_1, h_2 \in \mathsf{F}(X;\sigma,\delta)$:

- $\deg(h_1 h_2) = \deg h_1 + \deg h_2$
- $\deg(h_1 + h_2) \leqslant \deg h_1 + \deg h_2$

Theorem: Bound on quasideterminant degree

Let $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $\deg A_{ij} \leqslant d$. For all p, q such that $\det_{pq} A$ is defined, we have

$$-(n-1)d \leqslant \deg \det_{pq} A \leqslant n \deg A \qquad \text{or} \quad \det_{pq} A = 0$$

Proof

Use induction on the recursive formulation:

$$\det_{pq}(A) = A_{pq} - \sum_{i\neq p, j\neq q} A_{pi}(\det_{ji}(A^{(pq)}))^{-1} A_{jq}$$

Difficulty (but not really): not all quasideterminants are defined.

Implications

Corollary: Bound on inverse degree

Let $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $A_{ij} = 0$ or $0 \leqslant \deg A_{ij} \leqslant d$, and $B = A^{-1}$. Then $\deg B \leqslant n \deg A$.

Implications

Corollary: Bound on inverse degree

Let $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $A_{ij} = 0$ or $0 \leqslant \deg A_{ij} \leqslant d$, and $B = A^{-1}$. Then $\deg B \leqslant n \deg A$.

Hermite form degree bounds

$A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with Hermite form $H \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ and unimodular $U \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $UA = H$.

$$A \mapsto H = UA = \begin{pmatrix} H_{11} & * & \cdots & * \\ & H_{22} & \cdots & \vdots \\ & & \ddots & * \\ & & & H_{nn} \end{pmatrix}$$

Then $\sum \deg H_{ii} = \deg \delta\varepsilon\tau A \leqslant nd$, $\deg U \leqslant (n-1)\deg A$.

Implications

Corollary: Bound on inverse degree

Let $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $A_{ij} = 0$ or $0 \leqslant \deg A_{ij} \leqslant d$, and $B = A^{-1}$. Then $\deg B \leqslant n \deg A$.

Jacobson form degree bounds

$A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with Hermite form $H \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ and unimodular $U \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ with $UA = H$.

$$A \mapsto H = UAV = \begin{pmatrix} J_{11} & & & \\ & J_{22} & & \\ & & \ddots & \\ & & & J_{nn} \end{pmatrix}$$

Then $\sum \deg J_{ii} = \deg \delta\varepsilon\tau A \leqslant nd$, $\deg U, V \leqslant (n-1)\deg A$.

Quasideterminants and Dieudonné determinant

The Dieudonné determinant can be expressed in terms of quasideterminants:

For $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$:

$$\delta\varepsilon\tau(A) = \mathsf{det}_{11}(A)\cdot \mathsf{det}_{22}(A^{(11)})\cdots \mathsf{det}_{nn}(A^{(1\ldots n-1,1\ldots,n-1)})$$

and it easily follows that

$$\deg \delta\varepsilon\tau(A) \leqslant n\cdot \deg A$$

Also, if $U \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ is unimodular then $\deg \delta\varepsilon\tau\, U = 0$.

Linear Systems Method for Hermite Form Computation

Kaltofen et al. (1987), Storjohann (1994), Labhalla et al., (1996) reduce Hermite form of $A \in \mathsf{F}[x]^{n\times n}$ to solving $O(n^2 d) \times O(n^2 d)$ system of linear equations over F.

- Effective when $\mathsf{F} = \mathbb{Q}(t)$ and there is growth both in the degrees (in t) and the size of the coefficients in $\mathbb{Q}$.
  - The coefficients (in $\mathbb{Q}(t)$) are solutions to linear equations.
- The bounds on the sizes of entries tend to be tight, though the complexity is high (but polynomial in the input size).
- We will adapt this method to the non-commutative $\mathbb{Q}(t)[X;\delta]$, and more generally $\mathsf{F}[X;\sigma,\delta]$.

A pseudo-linear equation for entries in Hermite form

Given: $A \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ of full left row rank with $\deg A \leqslant d$

$(d_1,\ldots,d_n) \in \mathbb{N}^n$

Consider the system

$$PA = G$$

where $P, G \in \mathsf{F}[X;\sigma,\delta]^{n\times n}$ restricted as follows:

- $\deg P_{ij} \leqslant (n-1)d + \max_{1\leqslant i\leqslant n} d_i$.
- G is upper triangular
- Every diagonal entry of G is monic
- Degree of off-diagonal entries is less than the degree of the diagonal entry below it.
- The degree of the ith diagonal entry of G is d_i.

Theorem

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:

- *There exists at least one pair P, G with $PA = G$, as previously, if and only if $d_i \geqslant h_i$ for $1 \leqslant i \leqslant n$;*
- *If $d_i = h_i$ for $1 \leqslant i \leqslant n$ then G is the Hermite form of A and P is a unimodular matrix.*

This theorem allows us to perform binary search for the correct degree sequence.

The Linear Systems Method over $\mathsf{F}[X;\sigma,\delta]$

Express pseudo-linear system $PA = G$ as a linear system over F

$$\widehat{P}\,\widehat{A} = \widehat{G}$$

for

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

where $\beta = (n-1)d + \max_{1\leqslant i\leqslant n} d_i$. The entries of $\widehat{A}$ are obtained from A in such a way that:

- A_{ij} replaced by the $(\beta+1)\times(\mu+1)$ block where $\mu = \beta + d$.
- Its ℓth row is $(A^{[\ell]}_{ij\mu}, ..., A^{[\ell]}_{ij0})$ such that

$$X^{\ell} A_{ij} = A^{[\ell]}_{ij0} + \cdots + A^{[\ell]}_{ij\mu} X^{\mu}.$$

Similar to Li (1998) for Sylvester matrices.

The system is linear in indeterminates of $\widehat{P}$ and $\widehat{G}$, with $O(n^3 d)$ equations and $O(n^3 d)$ unknowns in F.

Can be reduced to $O(n^2 d)$, but that is probably “optimal”.

Linear Systems Method: Example

Back to $\mathsf{F} = \mathbb{Q}(t)[X;\delta]$

$$A = \begin{pmatrix} 2tX & t + (1+4t)X \\ 2t + tX & 9t + (1+5t)X \end{pmatrix}$$

and given $\vec{d} = (0,1)$. Then $\beta = (n-1)d + \max_{1 \leqslant i \leqslant n} d_i = 2$. We want to show how A_{11} is expanded in $\widehat{A}$:

$$\widehat{A} \mapsto \left(\begin{array}{cccc|cccc} 0 & 2t & 0 & 0 & t & 1+4t & 0 & 0 \\ 0 & 2 & 2t & 0 & 1 & 4+t & 4t+1 & 0 \\ 0 & 0 & 4 & 2t & 0 & 2 & t+8 & 4t+1 \\ \hline 2t & t & 0 & 0 & 9t & 1+5t & 0 & 0 \\ 2 & 2t+1 & t & 0 & 9 & 9t+5 & 5t+1 & 0 \\ 0 & 4 & 2t+2 & t & 0 & 18 & 9t+6 & 5t+1 \end{array}\right) \in \mathbb{Q}[t]^{6\times 8}$$

$\widehat{P}$ and $\widehat{G}$ expand similarly, but we don't know all the coefficients

➡ Unknown coefficients satisfy linear equations over $\mathbb{Q}(t)$.

Summary Cost of Hermite Computation

Cost of Computing Hermite Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^β.

We can compute the Hermite form $H \in \mathbb{Q}(t)[X;\delta]^{n\times n}$ of A and a unimodular $U \in \mathbb{Q}(t)[X;\delta]^{n\times n}$ such that $UA = H$ with $O((n^7 d^3 + n^4 d^2 e)\beta^2 \log(nd))$ bit operations.

Summary Cost of Hermite Computation

Cost of Computing Hermite Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^β.

We can compute the Hermite form $H \in \mathbb{Q}(t)[X;\delta]^{n\times n}$ of A and a unimodular $U \in \mathbb{Q}(t)[X;\delta]^{n\times n}$ such that $UA = H$
in polynomial time.

Coefficients of entries of U, H have $O(n^2 d\beta \log(nd))$ bits.

From Hermite to Jacobson

Focus on differential polynomials: $\mathbb{Q}(t)[X;\delta]$

Idea: a random conditioning makes the diagonal of the Hermite form equal to the diagonal of the Jacobson form.

Theorem

Let $A \in \mathbb{Q}(t)[X;\delta]^{n\times n}$. Let $V \in \mathbb{Q}[t]^{n\times n}$ be lower triangular with 1's on the diagonal, and subdiagonal "randomly" from $\mathbb{Q}(t)$. With high probability the Hermite form of AV has shape

$$\begin{pmatrix} 1 & 0 & \dots & \dots & * \\ & 1 & 0 & \dots & * \\ & & \ddots & \ddots & \vdots \\ & & & 1 & * \\ & & & & \varphi \end{pmatrix} \in \mathbb{Q}(t)[X;\delta]$$

From Hermite to Jacobson

Focus on differential polynomials: $\mathbb{Q}(t)[X;\delta]$

Idea: a random conditioning makes the diagonal of the Hermite form equal to the diagonal of the Jacobson form.

Theorem

Let $A \in \mathbb{Q}(t)[X;\delta]^{n\times n}$. Let $V \in \mathbb{Q}[t]^{n\times n}$ be lower triangular with 1's on the diagonal, and subdiagonal "randomly" from $\mathbb{Q}(t)$. With high probability the Hermite form of AV has shape

$$\begin{pmatrix} 1 & 0 & \dots & \dots & * \\ & 1 & 0 & \dots & * \\ & & \ddots & \ddots & \vdots \\ & & & 1 & * \\ & & & & \varphi \end{pmatrix} \in \mathbb{Q}(t)[X;\delta]$$

Corollary

The Jacobson normal form of A is diag$(1,\dots,1,\varphi)$

From Hermite to Jacobson

Focus on differential polynomials: $\mathbb{Q}(t)[X;\delta]$

Idea: a random conditioning makes the diagonal of the Hermite form equal to the diagonal of the Jacobson form.

Theorem

Let $A \in \mathbb{Q}(t)[X;\delta]^{n\times n}$. Let $V \in \mathbb{Q}[t]^{n\times n}$ be lower triangular with 1's on the diagonal, and subdiagonal "randomly" from $\mathbb{Q}(t)$. With high probability the Hermite form of AV has shape

$$\begin{pmatrix} 1 & 0 & \dots & \dots & * \\ & 1 & 0 & \dots & * \\ & & \ddots & \ddots & \vdots \\ & & & 1 & * \\ & & & & \varphi \end{pmatrix} \in \mathbb{Q}(t)[X;\delta]$$

What is "randomly"?

Subdiagonal entries are chosen from $\mathbb{Z}[t]$ with degree at most nd and coefficients from $\{0,\dots,2nd\}$.

From Hermite to Jacobson

Focus on differential polynomials: $\mathbb{Q}(t)[X;\delta]$

Idea: a random conditioning makes the diagonal of the Hermite form equal to the diagonal of the Jacobson form.

Theorem

Let $A \in \mathbb{Q}(t)[X;\delta]^{n\times n}$. Let $V \in \mathbb{Q}[t]^{n\times n}$ be lower triangular with 1's on the diagonal, and subdiagonal "randomly" from $\mathbb{Q}(t)$. With high probability the Hermite form of AV has shape

$$\begin{pmatrix} 1 & 0 & \dots & \dots & * \\ & 1 & 0 & \dots & * \\ & & \ddots & \ddots & \vdots \\ & & & 1 & * \\ & & & & \varphi \end{pmatrix} \in \mathbb{Q}(t)[X;\delta]$$

Caveat

An inflation of the degree nd in t is substantial, and the randomization tends to destroy any nice structure.

Jacobson form via Hermite form

Let F $= \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Jacobson form via Hermite form

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1 + (t+2)X + X^2 & 2 + (2t+1)X & 1 + (1+t)X \\ 2t + t^2 + tX & 2 + 2t + 2t^2 + X & 4t + t^2 \\ 3 + t + (3+t)X + X^2 & 8 + 4t + (5+3t)X + X^2 & 7 + 8t + (2+4t)X \end{bmatrix}$$

Choose a random $V \in \mathbb{Z}[t]^{3\times 3}$

Our bounds say entries in V should be polynomials in $\mathbb{Z}[t]$, random coefficients from $\{0, ..., 11\}$, of degree 6.

Jacobson form via Hermite form

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Choose a random $V \in \mathbb{Z}[t]^{3\times 3}$

So let's "randomly" try $V = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

Jacobson form via Hermite form

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \left[\begin{array}{ccc} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{array}\right]$$

Precondition A

$$A \mapsto AV = \begin{pmatrix} \frac{8t^2+7t-2}{2t} + \frac{2t-1}{2t}X & 2t+1 & \frac{t+2t^2-2}{2t} - \frac{1}{2t} * X^2 \\ \frac{9}{2}t+3+\frac{3}{2}X & (t+2)+X & (1+\frac{7}{2}t)+1/2X \\ \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 & 0 & \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 \end{pmatrix}$$

Jacobson form via Hermite form

Let $\mathsf{F} = \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Precondition A

$$A \mapsto AV = \begin{pmatrix} \frac{8t^2+7t-2}{2t} + \frac{2t-1}{2t}X & 2t+1 & \frac{t+2t^2-2}{2t} - \frac{1}{2t} * X^2 \\ \frac{9}{2}t + 3 + \frac{3}{2}X & (t+2)+X & (1+\frac{7}{2}t) + 1/2X \\ \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 & 0 & \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 \end{pmatrix}$$

Hermite form

$$H = \begin{bmatrix} 1 & 0 & \text{[redacted]} + \text{[redacted]}X + \text{[redacted]}X^2 + \text{[redacted]}X^3 \\ 0 & 1 & \text{[redacted]} + \text{[redacted]}X + \text{[redacted]}X^2 + \text{[redacted]}X^3 \\ 0 & 0 & \begin{array}{l} -\frac{372t^7+1884t^6+2109t^5-1415t^4+492t^3+1044t^2-757t+63}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1} + \\ \quad \frac{60t^9+408t^8+795t^7-1060t^6-3847t^5+648t^4+497t^3-1400t^2+783t-52}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1}X \\ \quad +3\frac{60t^8+288t^7+219t^6-654t^5-249t^4+199t^3-220t^2+113t-4}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1}X^2 \\ \quad +\frac{\left(180t^6+504t^5-171t^4-432t^3+60t^2-156t+95\right)t}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1}X^3 + X^4 \end{array} \end{bmatrix}$$

Jacobson form via Hermite form

Let F $= \mathbb{Q}(t)$ and $A \in \mathbb{Q}(t)[X;\delta]^{3\times 3}$, where $Xt = tX + 1$.

$$A = \begin{bmatrix} 1+(t+2)X+X^2 & 2+(2t+1)X & 1+(1+t)X \\ 2t+t^2+tX & 2+2t+2t^2+X & 4t+t^2 \\ 3+t+(3+t)X+X^2 & 8+4t+(5+3t)X+X^2 & 7+8t+(2+4t)X \end{bmatrix}$$

Precondition A

$$A \mapsto AV = \begin{pmatrix} \frac{8t^2+7t-2}{2t} + \frac{2t-1}{2t}X & 2t+1 & \frac{t+2t^2-2}{2t} - \frac{1}{2t} * X^2 \\ \frac{9}{2}t+3+\frac{3}{2}X & (t+2)+X & (1+\frac{7}{2}t)+1/2X \\ \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 & 0 & \frac{-2}{t} + \frac{t^2-1+2t}{t}X + X^2 \end{pmatrix}$$

Jacobson form

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \begin{array}{l} -\frac{372t^7+1884t^6+2109t^5-1415t^4+492t^3+1044t^2-757t+63}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1} + \\ \quad \frac{60t^9+408t^8+795t^7-1060t^6-3847t^5+648t^4+497t^3-1400t^2+783t-52}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1} X \\ \quad +3\frac{60t^8+288t^7+219t^6-654t^5-249t^4+199t^3-220t^2+113t-4}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1} X^2 \\ \quad + \frac{\left(180t^6+504t^5-171t^4-432t^3+60t^2-156t+95\right)t}{60t^6+48t^5-33t^4+2t^3-28t^2+6t+1} X^3 + X^4 \end{array} \end{bmatrix}$$

Why randomization works

Theorem

Let $f, g \in \mathbb{Q}(t)[X; \delta]$*. Then for a random* $w \in \mathsf{F}[t]$ *of degree* $\max\{\deg_X g, \deg_X g\}$*, we have* $\gcd(fw, g) = 1$*.*

Proof.

- Show that for any f, g there exists a w of degree $\max\{\deg_X g, \deg_X g\}$ such that $\gcd(fw, g) = 1$.
- Use a non-commutative Sylvester-like resultant to show that for this works almost all w.

$\square$

Complexity

Cost of Computing Jacobson Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^{β}.

Cost to compute J, U, $V \in \mathbb{Q}(t)[X;\delta]^{n\times n}$:

Complexity

Cost of Computing Jacobson Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^β.

Cost to compute J, U, $V \in \mathbb{Q}(t)[X;\delta]^{n\times n}$: **Polynomial time.**

Complexity

Cost of Computing Jacobson Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^β.

Cost to compute J, U, $V \in \mathbb{Q}(t)[X;\delta]^{n\times n}$:
$O((n^8 d^4 + n^5 d^3 e)\beta^2 \log(nd))$ bit operations. **Oooph.**

Complexity

Cost of Computing Jacobson Form over $\mathbb{Q}(t)[X;\delta]$

Let $A \in \mathbb{Z}[t][X;\delta]^{n\times n}$ with $\deg_X A \leqslant d$, $\deg_t A \leqslant e$, and coefficients of A_{ij} have absolute value at most 2^β.

Cost to compute J, U, $V \in \mathbb{Q}(t)[X;\delta]^{n\times n}$:
$O((n^8 d^4 + n^5 d^3 e)\beta^2 \log(nd))$ bit operations. **Oooph.**

Excuse: output is probably pretty big:

- U is $n \times n$ of degree nd in $\mathcal{D}$ and $(n-1)/n \cdot n^2 de$ in t.
- Total output size: $O(n^5 d^2 e)$ elements of $\mathbb{Q}$
- Coefficients in $\mathbb{Q}$ seem to have $\gg n^2 d\beta$ bits each
- Not really proven but we suspect it...

Conclusion

Have algorithms for Hermite and Jacobson form of a matrix over $\mathsf{F}[X;\sigma,\delta]$ which requires polynomial in the input size, accounting for *all coefficient and degree growth*.

Future work

- "Beautification" of Jacobson form
- Faster algorithms for Hermit/Jacobson form in $\mathsf{F}[X;\sigma,\delta]$. Algorithms over $\mathsf{F}[x]$ are still much faster, and there is no particularly good reason for this.
- Probably via a faster method for Popov form computation.
- Use the bounds provided by the linear systems method to allow for "modular" methods with Khochtali, Rosenkilde, Storjohann (2017)

References

- M. Giesbrecht and M. Sub Kim. *Computation of the Hermite form of a Matrix of Ore Polynomials*. Journal of Algebra, v. 376, pp. 341–362, 2013.
- M. Giesbrecht and A. Heinle. *A polynomial-time algorithm for the Jacobson form of a matrix of Ore polynomials*. Proc. CASC 2012, pp. 117-128. Lecture Notes in Computer Science Volume, v.7442.
- A. Heinle, *Computational Approaches to Problems in Noncommutative Algebra – Theory, Applications and Implementations*. PhD Thesis, U. Waterloo. 2017.
- A. Heinle, Factorization, *Similarity and Matrix Normal Forms over certain Ore Domains*. Master's Thesis, RWTH Aachen. 2012