Normalization of Row Reduced Matrices

Soumojit Sarkar
soumojitsarkar@gmail.com

Arne Storjohann
astorjoh@uwaterloo.ca

David R. Cheriton School of Computer Science
University of Waterloo, Ontario, Canada N2L 3G1

ABSTRACT

This paper gives gives a deterministic algorithm to transform a row reduced matrix to canonical Popov form. Given as input a row reduced matrix R over $\mathsf{K}[x]$, K a field, our algorithm computes the Popov form in about the same time as required to multiply together over $\mathsf{K}[x]$ two matrices of the same dimension and degree as R. We also show that the problem of transforming a row reduced matrix to Popov form is at least as hard as polynomial matrix multiplication.

Categories and Subject Descriptors

G.4 [**Mathematical Software**]: Algorithm Design and Analysis; I.1.2 [**Symbolic and Algebraic Manipulation**]: Algorithms; F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems

General Terms

Algorithms

Keywords

Popov form, Polynomial matrices

1. INTRODUCTION

This paper considers the problem of lattice reduction, or row reduction, for matrices over the ring $\mathsf{K}[x]$ of univariate polynomials with coefficients from a field K. Row reduction of a matrix A over $\mathsf{K}[x]$ is the problem of finding a basis with row degrees as small as possible for the lattice $\mathcal{L}(A)$ generated by all $\mathsf{K}[x]$-linear combinations of rows of A. For the following example, recall that a matrix $U \in \mathsf{K}[x]^{n\times n}$ is unimodular precisely when $\det U$ is a nonzero constant from K. Two matrices $A, R \in \mathsf{K}[x]^{n\times n}$ are *left equivalent* (i.e., the rows of A and R generate the same lattice) if and only if $A = UR$ for $U \in \mathsf{K}[x]^{n\times n}$ a unimodular matrix. We remark that in the literature some authors (for example [4]) prefer to consider the equivalent but transposed situation of column reduction, where the unimodular transform on the right.

ISSAC'11, June 8–11, 2011, San Jose, California, USA.
Copyright 2011 ACM 978-1-4503-0675-1/11/06 ...$10.00.

Example 1. *Let us indicate a polynomial of degree t with $[t]$. The following shows the degree structure in a particular matrix $A \in \mathsf{K}[x]^{4\times 4}$, a row reduced form R of A, and the unimodular matrix U such that $A = UR$.*

$$A = \begin{bmatrix} [13] & [13] & [12] & [12] \\ [13] & [13] & [12] & [12] \\ [13] & [13] & [12] & [12] \\ [13] & [13] & [12] & [12] \end{bmatrix}$$

$$= \overset{U}{\begin{bmatrix} [12] & [11] & [11] & [9] \\ [12] & [11] & [11] & [9] \\ [12] & [11] & [11] & [9] \\ [12] & [11] & [11] & [9] \end{bmatrix}} \overset{R}{\begin{bmatrix} [1] & [1] & [1] & [1] \\ [2] & [2] & [2] & [2] \\ [2] & [2] & [2] & [2] \\ [4] & [4] & [4] & [4] \end{bmatrix}}$$

Let $A \in \mathsf{K}[x]^{n\times n}$ be nonsingular. A fast Las Vegas probabilistic algorithm for computing a reduced basis R of A is given in [6]. Our main contribution in this paper is a deterministic algorithm that computes the canonical Popov reduced basis P, together with the unimodular matrix U such that $A = UP$, in about the same time as required to multiply together two polynomial matrices of the same dimension and degree as A. To clearly state our contributions, and to compare with previous work, we recall from [8, page 385] the precise definition of a row reduced form and the normalization conditions required for a row reduced form to be in canonical Popov form.

Let v be a row vector over $\mathsf{K}[x]$. The degree of v, denoted by $\deg v$, is the maximal degree of all entries. The *pivot index* of v, denoted by $\mathrm{piv}(v)$ is the index of the rightmost entry of degree $\deg v$. The *leading coefficient* vector $\mathrm{LC}(v)$ over K is obtained by taking the coefficient of $x^{\deg v}$ of all entries of v. Let A be a matrix over $\mathsf{K}[x]$. The degree of A, denoted by $\deg A$, is the maximal degree of its rows. The leading coefficient matrix of A, denoted by $\mathrm{LC}(A)$, is the matrix over K formed by taking the leading coefficient of each row of A.

Definition 2. *A nonsingular matrix*

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} = \begin{bmatrix} \vec{p}_1 \\ \hline \vec{p}_2 \\ \hline \vdots \\ \hline \vec{p}_n \end{bmatrix} \in \mathsf{K}[x]^{n\times n}$$

is row reduced *if $\mathrm{LC}(P)$ is nonsingular. If, in addition, P satisfies the following normalization conditions it is in* Popov form.

(i) *The pivot indices $\mathrm{piv}(\vec{p}_1), \ldots, \mathrm{piv}(\vec{p}_n)$ are distinct.*

(ii) *The pivot entries* $p_{1,\mathrm{piv}(\vec{p}_1)}, \ldots, p_{n,\mathrm{piv}(\vec{p}_n)}$ *are monic.*

(iii) $\deg \vec{p}_i \le \deg \vec{p}_{i+1}$ *for* $1 \le i < n$, *and if* $\deg \vec{p}_i = \deg \vec{p}_{i+1}$ *then* $\mathrm{piv}(\vec{p}_i) < \mathrm{piv}(\vec{p}_{i+1})$.

(iv) *Nonpivot entries have degree less than that of the pivot entry in the same column.*

If P *satisfies only condition (i) it is said to be in* weak Popov form *[9].*

Any nonsingular $A \in \mathsf{K}[x]^{n \times n}$ has a unique decomposition $A = UP$ with U unimodular and P in Popov form. The Popov form is a canonical form for left equivalence which has row degrees as small as possible, in particular, $\deg P \le \deg A$. We also remark that the multi-sets of row degrees of row reduced forms that are left equivalent are identical.

Example 3. *Consider the row reduced form* R *from Example 1. The following shows the possible degree structure in a weak Popov form* W *of* R*, and in the canonical Popov form* P *of* R*. The pivot entries in each row have been underlined.*

$$\overset{R}{\begin{bmatrix} [1] & [1] & [1] & \underline{[1]} \\ [2] & [2] & [2] & \underline{[2]} \\ [2] & [2] & [2] & \underline{[2]} \\ [4] & [4] & [4] & \underline{[4]} \end{bmatrix}} \rightarrow \overset{W}{\begin{bmatrix} [1] & [1] & [1] & \underline{[1]} \\ [1] & [2] & \underline{[2]} & [1] \\ \underline{[2]} & [1] & [1] & [1] \\ [3] & \underline{[4]} & [3] & [3] \end{bmatrix}} \rightarrow \overset{P}{\begin{bmatrix} [1] & [1] & [1] & \underline{[1]} \\ \underline{[2]} & [1] & [1] & [0] \\ [1] & [2] & \underline{[2]} & [0] \\ [1] & \underline{[4]} & [1] & [0] \end{bmatrix}}$$

Algorithms and complexity analysis for computing row reduced forms of matrices over $\mathsf{K}[x]$ are given in [9, 6, 10, 4], see also the references in [10]. In this paper, cost estimates will be given in terms of field operations from K, and we use ω for the exponent of matrix multiplication: two $n \times n$ matrices over a commutative ring can be multiplied in $O(n^{\omega})$ operations from the ring.

Let $A \in \mathsf{K}[x]^{n \times n}$ be nonsingular with $\deg A = d$. The deterministic algorithm in [9] computes the Popov form P of A in time $O(n^3 d^2)$. The algorithm in [9] is inherently iterative and does not seem amenable to a recursive approach which might introduce fast matrix and polynomial arithmetic. In [6] a Las Vegas randomized algorithm is given to compute a row reduced form of A with expected running time $O\tilde{~}(n^{\omega} d)$, which is about the same time as required to multiply together two polynomial matrices of the same dimension and degree as A. Our first contribution in this paper is to give an $O\tilde{~}(n^{\omega} d)$ deterministic algorithm to transform a row reduced matrix (such as produced by the algorithm in [6]) to Popov form. To the best of our knowledge, a transformation from row reduced from to Popov form in this time bound was not previously know. Note that in the particular case when all rows of a row reduced form R are equal, we can transform R to Popov form P in time $O(n^{\omega}\, d)$ using the identity $P = \mathrm{LC}(R)^{-1} R$. Our effort in this paper is devoted to the more subtle case when the row degrees of R are distinct.

On the one hand, for many applications a non canonical row reduced form R of A will suffice. In particular, a row reduced form gives a basis for $\mathcal{L}(A)$ that has row degrees as small as possible, and will satisfy the highly useful *predictable degree* property [8]: for polynomials $u_1, \ldots, u_n \in \mathsf{K}[x]$, we have $\deg u_1 \vec{p}_1 + \cdots + u_n \vec{p}_n = \max_i \{\deg u_i + \deg \vec{p}_i\}$.

On the other hand, computing the Popov form has some obvious advantages. Being canonical, equality of two lattices over $\mathsf{K}[x]$ can be determined by checking that their Popov basis are identical. If asked for a basis for a lattice over $\mathsf{K}[x]$, returning the Popov instead of only a row reduced form is analogous to a computer algebra system returning the normalized (i.e., monic) gcd of two scalar polynomials. Indeed, given two nonsingular matrices $A, B \in \mathsf{K}[x]^{n \times n}$, the Popov basis P of the lattice generated by the rows of A and B gives a canonical matrix greatest common right divisor of A and B: A and B can be expressed as $A = U_1 P$ and $B = U_2 P$ for polynomial matrices U_1 and U_2 for which there exists polynomial matrices V_1 and V_2 such that $V_1 U_1 + V_2 U_2 = I_n$.

To illustrate the analogy between the Popov form and the normalized monic gcd, it is useful to consider the definition of Popov form used in [4], which, up to a (unique) row permutation, is identical to the classical one we have given in Definition 2: condition (iii) is replaced with the condition that $\mathrm{piv}(\vec{p})_i = i$, that is, the rows are permuted so that the pivots are on the diagonal. Following [4, Definition 2.1], a row reduced matrix P as in (1) is in Popov form precisely when $\mathrm{LC}(P)$ is lower triangular and the normalization condition $\mathrm{LC}(P^T) = I_n$ is satisfied. Given the Popov form P of A, we can exploit the normalization condition $\mathrm{LC}(P^T) = I_n$ to get a fast algorithm that computes $U = AP^{-1}$ deterministically.

Producing a canonical form is also advantageous from an algorithmic point of view: a randomized Las Vegas algorithm for computing the Popov form P, instead of an arbitrary row reduced form R, will always return the same result even if different random choices are made. Many randomized algorithms require that the field K be large enough to ensure a positive probability of success. For example, the algorithm for row reduction in [6] first performs a random shift of variable $x \to x - \gamma$ to ensure that x does not divide $\det A$. To ensure a probability of success at least $1/2$ in the worst case, γ should be chosen form a subset of K of size at least $2nd$. If $\#\mathsf{K}$ is too small, a common technique is to work over a small algebraic extension $\bar{\mathsf{K}}$ of K that contains sufficiently many elements. However, a row reduced form R of $A \in \mathsf{K}[x]^{n \times n}$ may be over $\bar{\mathsf{K}}[x]$ if computed over $\bar{\mathsf{K}}[x]$. Nonetheless, even if we pass over an algebraic extension, the Popov form P must be over the ground field: $A \in \mathsf{K}[x]^{n \times n} \to \bar{R} \in \bar{\mathsf{K}}[x]^{n \times n} \to P \in \mathsf{K}[x]^{n \times n}$.

Our algorithm to transform R to P proceeds in two phases as illustrated in Example 3: first we transform R to a weak Popov form W, then we transform W to Popov form P. The first phase uses a careful modification of the `LUP` decomposition algorithm described in [1], and the second phase utilizes the fast minimal approximant basis algorithm of [6].

The rest of this paper is organized as follows. Section 2 recalls some facts about row reduced bases. Section 3 gives the algorithm to transform a row reduced form to weak Popov form. Section 4 gives an algorithm to go from weak Popov to Popov form. Section 5 gives the deterministic algorithm to produce the decomposition $A = UP$. Section 6 concludes, and offers a simple reduction of the problem of polynomial matrix multiplication to that of transforming a row reduced form to Popov form. Actually, we show that even the problem of transforming a matrix in weak Popov form to Popov form is as hard as polynomial matrix multiplication.

Cost model

Algorithms are analysed by bounding the number of required field operations from a field K on an algebraic random access machine; the operations $+$, $-$, $\times$ and "divide by a nonzero" involving two field elements have unit cost.

We use ω to denote the exponent of matrix multiplication: two $n \times n$ matrices over a ring R can be multiplied with $O(n^{\omega})$ ring operations from R. We use M for polynomial multiplication: let $\mathsf{M} : \mathbb{Z}_{\geq 0} \to \mathbb{R}_{>0}$ be such that polynomials in $\mathsf{K}[x]$ of degree bounded by d can be multiplied using at most $\mathsf{M}(d)$ field operations from K. We refer to [5] for more details and references about ω and M. We assume that $2 < \omega \leq 3$, and that $\mathsf{M}(ab) \leq \mathsf{M}(a)\mathsf{M}(b)$ for $a, b \in \mathbb{Z}_{>1}$. Some of our complexity estimates will explicitly make the assumption that $\mathsf{M}(t) \in O(n^{\omega-1})$. This assumption states that if fast matrix multiplication techniques are used, then fast polynomial multiplication should also be used.

Given two polynomials $a, b \in \mathsf{K}[x]$ with b nonzero, we denote by $\mathrm{Rem}(a, b)$ and $\mathrm{Quo}(a, b)$ the unique polynomials such that $a = \mathrm{Quo}(a, b)\, b + \mathrm{Rem}(a, b)$ with $\deg \mathrm{Rem}(a, b) < \deg b$. If a and b have degree bounded by d then both the Rem and Quo operation have cost $O(\mathsf{M}(d))$, and if b is a power of x both operations are free in our cost model. If the first argument of Rem or Quo is a matrix or vector the intention is to apply the function elementwise to the entries.

It will be useful to define an additional function B to bound the cost of the extended gcd operation, as well as other gcd-related computations. We can take either $\mathsf{B}(d) = \mathsf{M}(d) \log d$ or $\mathsf{B}(d) = d^2$. Then the extended gcd problem with two polynomials in $\mathsf{K}[x]$ of degree bounded by d can be solved in time $O(\mathsf{B}(d))$.

2. PRELIMINARIES

Row reduced and Popov forms are defined for matrices of arbitrary shape and rank profile. In this paper, we restrict ourselves to matrices of full row rank. The following definition generalizes Definition 2 to the case of full row rank matrices.

Definition 4. *A full row rank matrix*

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{bmatrix} = \begin{bmatrix} \vec{p}_1 \\ \hline \vec{p}_2 \\ \hline \vdots \\ \hline \vec{p}_n \end{bmatrix} \in \mathsf{K}[x]^{n \times m}$$

is row reduced *if* $\mathrm{LC}(P)$ *has full row rank n. If, in addition, P satisfies the following normalization conditions then it is in* Popov form.

(i) The pivot indices $\mathrm{piv}(\vec{p}_1), \ldots, \mathrm{piv}(\vec{p}_n)$ *are distinct.*

(ii) The pivot entries $p_{1,\mathrm{piv}(\vec{p}_1)}, \ldots, p_{n,\mathrm{piv}(\vec{p}_n)}$ *are monic.*

(iii) $\deg \vec{p}_i \leq \deg \vec{p}_{i+1}$ *for* $1 \leq i < n$, *and if* $\deg \vec{p}_i = \deg \vec{p}_{i+1}$ *then* $\mathrm{piv}(\vec{p}_i) < \mathrm{piv}(\vec{p}_{i+1})$.

(iv) $\deg p_{k,\mathrm{piv}(\vec{p}_i)} < \deg p_{i,\mathrm{piv}(\vec{p}_i)}$ *for* $k \in \{1, 2, \ldots, i-1, i+1, i+2, \ldots, n\}$, $1 \leq i \leq n$.

If P satisfies only condition (i) it is said to be in weak Popov form *[9].*

The following lemma recalls an essential feature of row reduced bases.

Lemma 5. *[8, Theorem 6.3-13] If* $R \in \mathsf{K}[x]^{n \times m}$ *is row reduced and* $v = \begin{bmatrix} v_1 \cdots v_n \end{bmatrix} \in \mathsf{K}[x]^{1 \times n}$, *then* $\deg vR = \max_i\{\deg v_i + \deg \mathrm{Row}(R, i)\}$.

In the following lemma, we use $\bar{*}$ to denote a square nonsingular matrix over K, and $*^d$ to denote a rectangular matrix over $\mathsf{K}[x]$ of degree bounded by d. The next lemma follows as a corollary of Lemma 5.

Lemma 6. *Let* $R, \bar{R} \in \mathsf{K}[x]^{n \times m}$ *be full row rank and row reduced matrices that are left equivalent. If both R and $\bar{R}$ have rows ordered such that degrees are nondecreasing, then the degrees of the rows of R and $\bar{R}$ are the same. Furthermore, if $d_1, d_2, \cdots, d_k$ is the nondecreasing sequence of distinct degrees of the rows of R, then*

$$\begin{array}{cccc} T & R & & \bar{R} \\ \begin{bmatrix} \overline{*} & & & \\ *^{d_2-d_1} & \overline{*} & & \\ \vdots & \vdots & \ddots & \\ *^{d_k-d_1} & *^{d_k-d_2} & \cdots & \overline{*} \end{bmatrix} & \begin{bmatrix} R^{[d_1]} \\ R^{[d_2]} \\ \vdots \\ R^{[d_k]} \end{bmatrix} & = & \begin{bmatrix} \bar{R}^{[d_1]} \\ \bar{R}^{[d_2]} \\ \vdots \\ \bar{R}^{[d_k]} \end{bmatrix} \end{array}$$

where the block decomposition is conformal, and $R^{[d_i]}$ denotes the submatrix of R comprised of the rows of degree d_i.

In the following corollary, let

$$X = \begin{bmatrix} x^{d_k-d_1}I & & & \\ & x^{d_k-d_2}I & & \\ & & \ddots & \\ & & & x^{d_k-d_k}I \end{bmatrix} \in \mathsf{K}[x]^{n \times n},$$

where the dimension of the diagonal block $x^{d_k-d_i}I$ corresponds to the row dimension of $R^{[d_i]}$, $1 \leq i \leq n$.

Corollary 7. *Let R, $\bar{R}$ and T be as in Lemma 6, and X be as in (1). Then* $L := \mathrm{LC}(x^{d_k}XTX^{-1}) \in \mathsf{K}^{n \times n}$, *with* $L\,\mathrm{LC}(R) = \mathrm{LC}(\bar{R})$.

Proof. The result can be seen most easily by passing over the ring of Laurent polynomials. Note that

$$(XTX^{-1})XR = X\bar{R},$$

with all rows in XR and $X\bar{R}$ of degree d_k, and $XTX^{-1} = L + O(x^{-1})_{x \to \infty}$ for $L \in \mathsf{K}^{n \times n}$. □

In the next section our goal is to find a matrix T as in Lemma 6 such that $W = TR \in \mathsf{K}[x]^{n \times n}$ is in weak Popov form. The following lemma, a corollary of Corollary 7, states that it is sufficient to solve this transformation to weak Popov form for a scalar input matrix, namely for $\mathrm{LC}(R) \in \mathsf{K}^{n \times n}$.

Lemma 8. *Let* $R \in \mathsf{K}[x]^{n \times m}$ *have full row rank, be row reduced, and have rows ordered so that degrees are nondecreasing. If* $\bar{T} \in \mathsf{K}^{n \times n}$ *is a unit lower triangular such that* $\bar{W} = \bar{T}\,\mathrm{LC}(R) \in \mathsf{K}^{n \times n}$ *is in weak Popov form, then* $T := X^{-1}\bar{T}X \in \mathsf{K}[x]^{n \times n}$ *is unimodular and* $W = TR \in \mathsf{K}[x]^{n \times n}$ *is in weak Popov form.*

Example 9. *The following partially specified matrix*

$$R = \begin{bmatrix} 73x + 56 & 68x + 24 & 65x + 90 & 3x + 16 \\ 78x^2 + \cdots & 59x^2 + \cdots & 69x^2 + \cdots & 3x^2 + \cdots \\ 60x^2 + \cdots & 41x^2 + \cdots & 83x^2 + \cdots & 5x^2 + \cdots \\ 75x^4 + \cdots & 94x^4 + \cdots & 70x^4 + \cdots & 3x^4 + \cdots \end{bmatrix}$$

is row reduced, where $\mathsf{K} = \mathbb{Z}/(97)$*. The following shows a transformation of* $\mathrm{LC}(R)$ *to weak Popov form* $\bar{W}$*.*

$$\overset{\bar{T}}{\begin{bmatrix} 1 & & & \\ 96 & 1 & & \\ 89 & 71 & 1 & \\ 3 & 38 & 33 & 1 \end{bmatrix}} \overset{\mathrm{LC}(R)}{\begin{bmatrix} 73 & 68 & 65 & 3 \\ 78 & 59 & 69 & 3 \\ 60 & 41 & 83 & 5 \\ 75 & 94 & 70 & 3 \end{bmatrix}} = \overset{\bar{W}}{\begin{bmatrix} 73 & 68 & 65 & 3 \\ 5 & 88 & 4 & \\ 67 & & & \\ & 3 & & \end{bmatrix}}$$

If we set

$$T = \overset{X^{-1}}{\begin{bmatrix} x^{-3} & & & \\ & x^{-2} & & \\ & & x^{-2} & \\ & & & 1 \end{bmatrix}} \overset{\bar{T}}{\begin{bmatrix} 1 & & & \\ 96 & 1 & & \\ 89 & 71 & 1 & \\ 3 & 38 & 33 & 1 \end{bmatrix}} \overset{X}{\begin{bmatrix} x^{3} & & & \\ & x^{2} & & \\ & & x^{2} & \\ & & & 1 \end{bmatrix}}$$
$$= \begin{bmatrix} 1 & & & \\ 96x & 1 & & \\ 89x & 71 & 1 & 0 \\ 3x^3 & 38x^2 & 33x^2 & 1 \end{bmatrix},$$

then $W = TR$ *is in weak Popov form with* $\bar{W} = \mathrm{LC}(W)$*.*

3. ROW REDUCED TO WEAK POPOV

Our goal is to transform a row reduced matrix to weak Popov form. By Lemma 8, it will be sufficient to handle the scalar case, that is, given a full row rank $R \in \mathsf{K}^{n\times m}$, compute a unit lower triangular transformation matrix $T \in \mathsf{K}^{n\times n}$ such that TR is in weak Popov form. Our approach is to compute a decomposition $R = LUP$ where L is unit lower triangular, U is upper triangular, and P is a permutation matrix. We accomplish this using a modification of the well known `LUP` decomposition algorithm described in [1, Page 236]. The following lemma gives the idea of our approach.

Lemma 10. *Let* $R \in \mathsf{K}^{n\times m}$ *have full row rank, and let* $R = LUP$ *be an* `LUP` *decomposition of* R*. If* $(p_1, \ldots, p_n)$ *is such that* p_i *is the index of integer* i *in the permuted tuple* $(1, 2, \ldots, m)P$*, then* $(UP)_{i,p_i}$ *is nonzero and entries in* UP *below* $(UP)_{i,p_i}$ *are zero,* $1 \le i \le n$*. Furthermore, if* $(UP)_{i,p_i}$ *is the rightmost nonzero entry in row* i *of* UP *for* $1 \le i \le n$*, then* $L^{-1}R = UP$ *is in weak Popov form.*

The following example is based on Example 9.

Example 11. *The following shows an* `LUP` *decomposition of a nonsingular* $R \in \mathbb{Z}_{97}^{4\times 4}$*.*

$$R = \overset{L}{\begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 34 & 26 & 1 & \\ 1 & 74 & 64 & 1 \end{bmatrix}} \overset{U}{\begin{bmatrix} 3 & 65 & 73 & 68 \\ & 4 & 5 & 88 \\ & & 67 & 0 \\ & & & 3 \end{bmatrix}} \overset{P}{\begin{bmatrix} & & & 1 \\ & & 1 & \\ 1 & & & \\ & 1 & & \end{bmatrix}}$$
$$= \begin{bmatrix} 73 & 68 & 65 & 3 \\ 78 & 59 & 69 & 3 \\ 60 & 41 & 83 & 5 \\ 75 & 84 & 70 & 3 \end{bmatrix}.$$

Now observe that $\bar{T}$ *and* $\bar{W}$ *in Example 9 are equal to* L^{-1} *and* UP*, respectively. But not every* `LUP` *decomposition leads to transformation to weak Popov form. For example,* R *has generic rank profile and so can be decomposed as the product of a unit lower triangular and upper triangular matrix.*

For $i = 1, 2, \ldots, n$, the iterative `LUP` decomposition algorithm chooses a nonzero pivot element in row i of the work matrix, postmultiplies the work matrix by a permutation P_i, swapping column i with a latter column, if needed, to ensure the pivot entry is located in column i, and zeroes out entries below the pivot entry by premultiplying the work matrix with a matrix L_i that is unit lower triangular with all entries zero except for possibly column i. Setting and $L := (L_n \cdots L_2 L_1)^{-1}$, $P := (P_1 P_2 \cdots P_n)^{-1}$ and U to be the final work matrix, gives an `LUP` decomposition. To ensure that the `LUP` decomposition produced will lead to a transformation to weak Popov form we need to specify how the pivot entries are chosen. Initialize a tuple $D = (1, 2, \ldots, n)$. After each row is processed the tuple D should be updated as $D := DP_i$. The pivot in row i is chosen to be the nonzero entry from among the last $n - i + 1$ entries of row i of the work matrix for which the corresponding component of D is maximal.

Example 12. *Let* R *be as in Example 11. Initialize* $D = (1, 2, 3, 4)$*. The first pivot we select is the right most element of the first row of* R*. This gives*

$$R_1 = \overset{L_1}{\begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ -34 & & 1 & \\ -1 & & & 1 \end{bmatrix}} \overset{R}{\begin{bmatrix} 73 & 68 & 65 & 3 \\ 78 & 59 & 69 & 3 \\ 60 & 41 & 83 & 5 \\ 75 & 84 & 70 & 3 \end{bmatrix}} \overset{P_1}{\begin{bmatrix} & & & 1 \\ & 1 & & \\ & & 1 & \\ 1 & & & \end{bmatrix}}$$
$$= \begin{bmatrix} 3 & 68 & 65 & 73 \\ & 88 & 4 & 5 \\ & 57 & 7 & 3 \\ & 16 & 5 & 2 \end{bmatrix}.$$

The updated D *is* $D = (4, 2, 3, 1)$*. The next pivot is thus chosen to be the third element of row 2 of* R_1*. The next elimination step gives*

$$R_2 = \overset{L_2}{\begin{bmatrix} 1 & & & \\ & 1 & & \\ & -26 & 1 & \\ & -74 & & 1 \end{bmatrix}} \overset{R_1}{\begin{bmatrix} 3 & 68 & 65 & 73 \\ & 88 & 4 & 5 \\ & 57 & 7 & 3 \\ & 16 & 5 & 2 \end{bmatrix}} \overset{P_2}{\begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix}}$$
$$= \begin{bmatrix} 3 & 65 & 68 & 73 \\ & 4 & 88 & 5 \\ & & & 67 \\ & & 3 & 20 \end{bmatrix}.$$

and D *is updated to* $D = (4, 3, 2, 1)$*.*

When applied to a full row rank $n \times m$ matrix, the base cases of the fast `LUP` decomposition algorithm will consist in computing an LUP decomposition of a nonzero $1 \times m$ matrix B which corresponds to the last m columns of a row of the work matrix, $1 \le m \le n$. By modifying the algorithm as follows, it will produce the same output as the iterative version with pivoting as specified above.

- Initialize $D = (1, 2, \ldots, n)$ at the start of the algorithm.
- At each base case involving a $B \in \mathsf{K}^{1\times m}$, compute the unique `LUP` decomposition $B = LUP$ which has P^{-1} equal to the permutation that interchanges column 1 and j, with j chosen so that $D[n - m + j]$ is maximal from among all j with $B[j]$ nonzero, $1 \le j \le m$. Update D by interchanging $D[n-m+1]$ and $D[n-m+j]$.

```
ReducedToWeakPopov($R, n, m, d$)

Input: A row reduced matrix $R \in \mathsf{K}[x]^{n\times m}$ with
       rank $n$ and $d = \deg R$.
Output: $W$, a weak Popov form of $R$.

  1. [Compute scalar transformation]
     Row permute $R$ so that degrees are nondecreasing.
     $\bar{R} := \mathrm{LC}(R)$;
     $L, U, P :=$ an LUP decomposition of $\bar{R}$ with
                  pivots chosen as described above;

  2. [Apply transformation]
     Let $d_i$ be the degree of row $i$ of $R$, $1 \le i \le n$.
     $X := \mathrm{Diag}(x^{d_1}, x^{d_2}, \ldots, x^{d_n})$;
     $\bar{T} := L^{-1}$;
     $W := X(\bar{T}(X^{-1}R))$;
     return $W$
```

Figure 1: Algorithm ReducedToWeakPopov

We obtain the following result as a corollary of Lemma 8 and [1, Theorem 6.4].

Theorem 13. *Algorithm* ReducedToWeakPopov *is correct. The cost of the algorithm is $O(mn^{\omega-1}\,d)$ operations from K.*

4. WEAK POPOV TO POPOV

In this section we show how to transform a full rank matrix $W \in \mathsf{K}[x]^{n\times m}$ that is in weak Popov form to Popov form. The following lemma observes that we can restrict our attention to the square nonsingular case.

Lemma 14. *Let $W \in \mathsf{K}[x]^{n\times m}$ have rank n and be in weak Popov form. If B is the submatrix of W comprised of the columns containing pivot entries, and $T \in \mathsf{K}[x]^{n\times n}$ is a unimodular matrix such that TB is in Popov form, then TW is the Popov form of W.*

Proof. Without loss of generality, up to a row permutation, assume W satisfies conditions (i) and (ii) of Definition 4. Then we can observe that the iterative algorithm of [9, Section 7] to transform W to Popov form P will maintain $\mathrm{piv}(\mathrm{Row}(W,i)) = \mathrm{piv}(\mathrm{Row}(P,i))$ for $1 \le i \le n$. Now, if $\bar{P}$ is the submatrix of P comprised of the columns containing the pivot entries, then $\bar{P}$ satisfies all conditions of Definition 2 and is in Popov form. Thus $\bar{P}$ is the Popov form TB of B. The result follows. □

The next lemma follows directly from Definition 2.

Lemma 15. *Let $P \in \mathsf{K}[x]^{n\times n}$ be nonsingular and in Popov form, and let c_i equal to the degree of the pivot entry in column i, $1 \le i \le n$. Set $X := \mathrm{Diag}(x^{d-c_1}, \ldots, x^{d-c_n})$, where $d = \deg P$. If Q is the permutation matrix such that QP has pivot entries located on the diagonal, then QPX is in Popov form with every row of degree d.*

Example 16. *The following shows the column shift of the Popov form from Example 1.*

$$
\begin{aligned}
QPX \;&=\; \overset{Q}{\begin{bmatrix} & 1 & & \\ & & & 1 \\ & & 1 & \\ 1 & & & \end{bmatrix}}
\overset{P}{\begin{bmatrix} [1] & [1] & [1] & \underline{[1]} \\ \underline{[2]} & [1] & [1] & [0] \\ [1] & [2] & \underline{[2]} & [0] \\ [1] & \underline{[4]} & [1] & [0] \end{bmatrix}}
\overset{X}{\begin{bmatrix} x^2 & & & \\ & 1 & & \\ & & x^2 & \\ & & & x^3 \end{bmatrix}} \\
&=\; \overset{QPX}{\begin{bmatrix} \underline{[4]} & [1] & [3] & [3] \\ [3] & \underline{[4]} & [3] & [3] \\ [3] & [2] & \underline{[4]} & [3] \\ [3] & [1] & [3] & \underline{[4]} \end{bmatrix}}.
\end{aligned}
$$

The next lemma follows from Definition 2.

Lemma 17. *If $R \in \mathsf{K}[x]^{n\times n}$ be a row reduced matrix with every row of degree d, then $\mathrm{LC}(R)^{-1}R$ is the Popov form of R and all its pivot elements are along the diagonal of the matrix.*

The following corollary of Lemmas 15 and 17 now shows how we may transform the problem of computing the Popov form of a weak Popov form to that of computing a row reduced basis of a suitably shifted matrix.

Theorem 18. *Let $B \in \mathsf{K}[x]^{n\times n}$ be nonsingular and in weak Popov form, and let c_i equal to the degree of the pivot entry in column i, $1 \le i \le n$. Let T be the unimodular matrix such that $P = TB$ is in Popov form, and let Q be the permutation matrix such that pivot entries in QP are on the diagonal. Set $d = \deg B$ and $X := \mathrm{Diag}(x^{d-c_1}, \ldots, x^{d-c_n})$. If $U \in \mathsf{K}[x]^{n\times n}$ is a unimodular matrix such that $R = UBX$ is row reduced, then $T := Q^{-1}\mathrm{LC}(UBX)^{-1}U \in \mathsf{K}[x]^{n\times n}$. Moreover, $\deg T \le d$.*

Proof. By Lemma 15 the matrix QPX will be in Popov form with all rows of degree d. Since QT is a unimodular matrix, $QP \equiv_{\mathrm{L}} B$ and so also $QPX \equiv_{\mathrm{L}} BX$. Since the Popov form QPX has all rows of degree d, the left equivalent reduced form UBX will also have all rows of degree d. Lemma 17 now shows that the following diagram commutes.

$$\begin{tikzcd}
B \arrow[r, "\text{Postmul. by } X"] \arrow[d, "\text{Premul. by } QT"'] & BX \arrow[d, "\text{Premul. by } \mathrm{LC}(UBX)^{-1}U"] \\
QP \arrow[r, "\text{Postmul. by } X"] & QPX
\end{tikzcd}$$

The claim that $T = Q^{-1}\mathrm{LC}(R)^{-1}U$ follows.

Now consider the degree of T. Since $P = TB$ is the Popov form of B, we have $\deg P \le \deg B = d$. The predictable degree property (Lemma 5) now implies that $\deg T \le d$. □

The final ingredient is the transformation of the matrix BX of Theorem 18 to row reduced form. To accomplish this we use a minimal approximant basis computation as described by [3, Theorem 5.2]. We will use algorithm PM-Basis of [6] to compute an order $3d+1$ minimal approximant $M \in \mathsf{K}[x]^{2n\times 2n}$ for the matrix

$$G = \left[\frac{BX}{-I_n}\right] \in \mathsf{K}[x]^{2n\times n}. \qquad (1)$$

Recall that M is a nonsingular row reduced matrix that gives a basis for the lattice $\{w \in \mathsf{K}[x]^{1\times n} \mid wG \equiv 0 \bmod x^{3d+1}\}$. We obtain the following result.

Lemma 19. *Let B and X be as in Theorem 18. If M is a minimal approximant basis of order $3d+1$ for G shown in (1), and $\left[\begin{array}{c|c}\bar{U} & \bar{R}\end{array}\right]$ is the submatrix of M comprised of the rows of degree bounded by d, with $\bar{U}$ of column dimension n, then $\bar{U}$ is unimodular and $\bar{R}$ is a row reduced form of BX.*

Proof. First note that the degree bounds $\deg \bar{U} \leq d$, $\deg \bar{R} \leq d$ and $\deg BX \leq 2d$, together with $\left[\begin{array}{c|c}\bar{U} & \bar{R}\end{array}\right] G \equiv 0 \bmod x^{3d+1}$, imply that

$$\left[\begin{array}{c|c}\bar{U} & \bar{R}\end{array}\right]\left[\begin{array}{c}BX\\ \hline -I_n\end{array}\right] = 0. \tag{2}$$

We will show in succession that the following hold:

(a) $\bar{U}$ has at most n rows.

(b) $\bar{U}$ is nonsingular.

(c) $\bar{U}$ is unimodular.

Using (c) together with (2) (i.e., $\bar{U}(BX) = \bar{R}$) shows that $\bar{R}$ is left equivalent to BX with all rows of $\bar{R}$ of degree d. Since the Popov form of BX has all rows of degree d, $\bar{R}$ must be a row reduced form of BX.

Claim (a): Since the rows of M are linearly independent, the row dimension of $\bar{U}$ can't be more than the dimension of the nullity of G, which is n.

Claim (b): From Theorem 18 we have $\left[\begin{array}{c|c}U & R\end{array}\right] G = 0$, with $\deg U, \deg R \leq d$. Since M is minimal approximant basis, all n linearly independent rows of $\left[\begin{array}{c|c}U & R\end{array}\right]$ must be generated by $\left[\begin{array}{c|c}\bar{U} & \bar{R}\end{array}\right]$. Since U is nonsingular and $\bar{U}$ has at most n rows, $\bar{U}$ must also be nonsingular.

Claim (c): From (2) we have $\bar{U}BX = \bar{R}$. Since $\bar{U}$ is nonsingular by claim (c), $\bar{R}$ is nonsingular also. The Popov form of BX has all rows of degree d, so $\deg\det BXnd$. Since $\deg \bar{R} \leq d$, we have $\deg\det \bar{R} \leq nd$. Finally, using $\bar{U}BX = \bar{R}$ gives that $\deg\det \bar{U} \leq \deg\det \bar{R} - \deg\det BX \leq 0$, showing that $\bar{U}$ is unimodular. □

Algorithm `WeakToPopov` is shown in Figure 2. By [6, Theorem 2.4], M is computed in $O(n^{\omega}\,\mathsf{B}(d))$ field operations from K. We obtain the following result.

Theorem 20. *Algorithm* `WeakToPopov` *is correct. The cost of the algorithm is $O(n^{\omega}\,\mathsf{B}(d) + mn^{\omega-1}\,\mathsf{M}(d))$ field operations from K.*

```
WeakToPopov($W, n, m, d$)

Input: A weak Popov form $W \in \mathsf{K}[x]^{n\times m}$ of rank $n$
        and degree $d$.
Output: $P$, the Popov form of $W$.

  1. [Extract pivot columns and scale]
     $B$ := submatrix of $W$ comprised of the columns
            containing pivot entries;
     Let $c_i$ be the degree of the pivot in column $i$ of $B$.
     $X := \mathrm{Diag}(x^{d-c_1}, x^{d-c_2}, \ldots, x^{d-c_n})$;

  2. [Minimal approximant basis computation]
     $G := \left[\begin{array}{c|c}BX & -I_n\end{array}\right]^T \in \mathsf{K}[x]^{2n\times n}$;
     $\delta := (0, \ldots, 0)$, of length $2n$;
     $M :=$ PM-Basis($G, 3d+1, \delta$);
     $\left[\begin{array}{c|c}U & R\end{array}\right]$ := the rows of $M$ that have
                  degree bounded by $d$;

  3. [Recover the Popov form of $W$]
     $T := \mathrm{LC}(R)^{-1}U$;
     $P := TW$;
     Permute rows of $P$ so that (iii) of Def. 4 holds;
     return $P$
```

Figure 2: Algorithm `WeakToPopov`

5. POPOV DECOMPOSITION OF NONSINGULAR MATRICES

Let $A \in \mathsf{K}[x]^{n\times n}$ be nonsingular of degree d. In this section we put together the results of the previous sections and give a deterministic algorithm to produce the decomposition $A = UP$ where P is the Popov form of A and U is unimodular.

Once the Popov form P has been computed, we can recover U as $U = AP^{-1}$. Let $X = \mathrm{Diag}(x^{c_1}, \ldots, x^{c_n})$, where c_i is the degree of the pivot entry in column i of P, $1 \leq i \leq n$. The Popov form P of A also satisfies $\deg P \leq \deg Ad$, and by Lemma 5, each row of U must also have degree bounded by d. Now note that

$$\begin{aligned} U &= AP^{-1}\\ &= (AX^{-1})(PX^{-1})^{-1}\\ &= \left(y^{-d}\,\overbrace{y^{d}(AX^{-1})|_{x=1/y}}^{D}\,{\overbrace{((PX^{-1})|_{x=1/y})}^{B}}^{-1} \right)|_{y=1/x} \end{aligned}$$

where D and B are over $\mathsf{K}[y]$. Since $\deg U \leq d$ and since $B(0)$ is invertible, we have $y^{d}\,U(y^{-1}) = \mathrm{Rem}(DB^{-1}, y^{d+1})$.

Algorithm `NonsingularPopovDecomp` shown in Figure 3 uses the scheme described above the compute U from P and A. Algorithm `RowReduce` used in phase 1 is described in [7]. `RowReduce` is a deterministic variant of the Las Vegas randomized algorithm for row reduction in [6] that, unlike the algorithm from [6], avoids the need to know *a a priori* or choose randomly an $\alpha \in \mathsf{K}$ such that $x - \alpha$ does not divide $\det A$. By [7, Theorem 36], the cost of computing R in phase 1 is $O(n^{\omega}(\log n)^2\,\mathsf{B}(d))$ field operations from K. The only computations in phase 2 requiring field operations is the computation of $\mathrm{Rem}(B^{-1}, y^{d+1})$ and the product $D\,\mathrm{Rem}(B^{-1}, y^{d+1})$. Since the constant coefficient of B is a permutation of I_n, the inverse of B up to order y^{d+1} can be computed using Newton iteration in time $O(n^{\omega}\,\mathsf{M}(d))$. We obtain the following result as a corollary of Theorems 13 and 20.

Theorem 21. *Algorithm* `NonsingularPopovDecomp` *is correct. The cost of the algorithm is $O(n^{\omega}(\log n)^2\,\mathsf{B}(d))$ field operations from K. This result assumes that $\mathsf{B}(t) \in O(t^{\omega-1})$.*

6. CONCLUSIONS AND FUTURE WORK

Given that the Popov form P has the same set of row degrees as a reduced form R, and only requires some ad-

```
NonsingularPopovDecomp($A, n, d$)

Input: A nonsingular matrix $A \in \mathsf{K}[x]^{n\times n}$ of degree $d$.
Output: $P, U \in \mathsf{K}[x]^{n\times n}$, with $P$ the Popov form of $A$
        and $A = UP$.

  1. [Compute the Popov form]
     $R :=$ RowReduce($A, n, d$);
     $W :=$ ReducedToWeakPopov($R, n, n, \deg R$);
     $P :=$ WeakToPopov($W, n, n, \deg W$);

  2. [Compute $U$]
     $X := \mathrm{Diag}(x^{c_1}, \ldots, x^{c_n})$, where $c_i = \deg \mathrm{Col}(P, i)$;
     $B := (PX^{-1})\,|_{x=1/y}$;
     $D := y^d(AX^{-1})\,|_{x=1/y}$;
     $E := y^{-d}\,\mathrm{Rem}(D\,\mathrm{Rem}(B^{-1}, y^{d+1}), y^{d+1})$;
     $U := E|_{y=1/x}$;
     return $P, U$
```

Figure 3: Algorithm `NonsingularPopovDecomp`

ditional normalization conditions to be satisfied, a natural question that arises is if the transformation from R to P is at least as hard as polynomial matrix multiplication: we answer this question affirmatively with a reduction similar to the well known reduction [1, Page 246] of scalar matrix multiplication to triangular matrix inversion.

Let $A, B \in \mathsf{K}[x]^{n\times n}$ have degree bounded by d. The following matrix C with degree bounded by $2d+1$ is row reduced since it is in weak Popov form:

$$C := \left[\begin{array}{c|c} x^{d+1}I_n & B \\ \hline -x^{d+1}A & x^{2d+1}I_n \end{array}\right] \in \mathsf{K}[x]^{2n\times 2n}.$$

The Popov form P of C is obtained as follows:

$$\left[\begin{array}{c|c} I & \\ \hline A & I \end{array}\right] \overset{C}{\left[\begin{array}{c|c} x^{d+1}I & B \\ \hline -x^{d+1}A & x^{2d+1}I \end{array}\right]} = \overset{P}{\left[\begin{array}{c|c} x^{d+1} & B \\ \hline & AB + x^{2d+1}I \end{array}\right]}.$$

We obtain the following result.

Theorem 22. *If we have an algorithm (algebraic* RAM*) for transforming a nonsingular $2n \times 2n$ row reduced matrix of degree $2d+1$ to Popov form with $P(n, d)$ operations from K, then two $n \times n$ matrices of degree d over $\mathsf{K}[x]$ can be multipied together with $P(n, d)$ operations from K.*

Our algorithms for transforming from row reduced to weak Popov, and from weak Popov to Popov, worked for rectangular input matrices of full row rank. Currently, our deterministic algorithm for computing the Popov decomposition requires the input matrix to be square and nonsingular. Randomization can be used to extend the algorithm to matrices of arbitrary shape and rank, but our ultimate goal is to obtain a deterministic algorithm for the general case.

7. REFERENCES

[1] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms.* Addison-Wesley, 1974.

[2] B. Beckermann and G. Labahn. A uniform approach for the fast computation of matrix–type Padé approximants. *SIAM Journal on Matrix Analysis and Applications*, 15(3):804–823, 1994.

[3] B. Beckermann, G. Labahn, and G. Villard. Shifted normal forms of polynomial matrices. In S. Dooley, editor, *Proc. Int'l. Symp. on Symbolic and Algebraic Computation: ISSAC '99*, pages 189—196. ACM Press, New York, 1999.

[4] B. Beckermann, G. Labahn, and G. Villard. Normal forms for general polynomial matrices. *Journal of Symbolic Computation*, 41(6):708–737, 2006.

[5] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra.* Cambridge University Press, 2 edition, 2003.

[6] P. Giorgi, C.-P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In R. Sendra, editor, *Proc. Int'l. Symp. on Symbolic and Algebraic Computation: ISSAC '03*, pages 135–142. ACM Press, New York, 2003.

[7] S. Gupta, S. Sarkar, A. Storjohann, and J. Valeriote. Triangular x-basis decompositions and derandomization of linear algebra algorithms over $\mathsf{K}[x]$. *Journal of Symbolic Computation.* Accepted for publication.

[8] T. Kailath. *Linear Systems.* Prentice Hall, Englewood Cliffs, N.J., 1980.

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

[10] G. Villard. Computing Popov and Hermite forms of polynomial matrices. In Y. N. Lakshman, editor, *Proc. Int'l. Symp. on Symbolic and Algebraic Computation: ISSAC '96*, pages 251–258. ACM Press, New York, 1996.