Fast Parallel Computation of Hermite and Smith Forms of Polynomial Matrices*

Erich Kaltofen
M. S. Krishnamoorthy

Rensselaer Polytechnic Institute
Department of Computer Science
Troy, New York 12181

B. David Saunders

University of Delaware
Department of Computer and Information Sciences
Newark, Delaware 19716

Abstract. Boolean circuits of polynomial size and poly-logarithmic depth are given for computing the Hermite and Smith normal forms of polynomial matrices over finite fields and the field of rational numbers. The circuits for the Smith normal form computation are probabilistic ones and also determine very efficient sequential algorithms. Furthermore, we give a polynomial-time deterministic sequential algorithm for the Smith normal form over the rationals. The Smith normal form algorithms are applied to the Rational canonical form of matrices over finite fields and the field of rational numbers.

Keywords: Parallel algorithm, Hermite normal form, Smith normal form, polynomial-time complexity.

1. Introduction

The main results of this paper establish fast parallel algorithms for computing the Hermite and Smith normal form of matrices with polynomial entries. The Hermite or Smith normal form of a square matrix is generally defined for the case of entries from a principal ideal domain. For example the entry domain may be the integers or univariate polynomials over a field. The forms are, roughly speaking, a triangularization, respectively a diagonalization, of the input matrix and they are computed entirely within the domain of the entries. Sequential algorithms for computing the forms are known at least since Hermite [7] and Smith [20], but it requires some effort to show that the forms can be computed in polynomial-time. We refer to Kannan and Bachem [13] for integer entries and Kannan [12] for polynomial entries. Applications of both forms include solving linear systems over the domain of entries, computing the geometric multiplicities of the

* This material is based upon work supported by the National Science Foundation under Grant No. MCS-83-14600. This paper appears in the SIAM J. on Algebraic and Discrete Methods, vol. 8, pp. 683--690 (1987).

eigenvalues of a matrix, computing the invariant factors of a matrix over a field, and others. For discussion of applications see [1] and [18].

We will show that computing the Hermite normal form over $F[x]$, F a field, is $\mathbf{NC}^1$ reducible to solving singular linear systems. We refer to Cook [4] for the definitions of the complexity classes **NC** and **RNC** and $\mathbf{NC}^1$ reductions. Since the class **NC** requires us to perform field operations on Boolean circuits, the previous claim is precise only for concrete fields such as **Q** or GF(p), the field with p elements. As a corollary we get from the parallel complexity of linear systems [2] and [16] that *HERMITE FORM* over $\mathbf{Q}[x]$ and GF(p)[x] is in $\mathbf{NC}^2$, where *HERMITE FORM* over D is the problem of computing Hermite normal forms over D. Our parallel reduction is completely different from any of the sequential solutions, discussed for example in [13]. Of course, it has Kannan's result that *HERMITE FORM* over $\mathbf{Q}[x]$ is in **P** as a consequence, where **P** is the class of sequential polynomial-time problems.

Secondly, we will present a probabilistic parallel algorithm for computing the Smith normal form over $F[x]$, that is we establish that *SMITH FORM* over $F[x]$ is in $\mathbf{RNC}^2$. The nature of our probabilistic algorithm is such that with controllably small probability an incorrect result might be returned, similar to the fast probabilistic parallel rank algorithm [2]. Since Kannan [12] does not prove that his sequential algorithm for *SMITH FORM* over $\mathbf{Q}[x]$ runs in polynomial-time we will also present another sequential algorithm with which we can establishing that *SMITH FORM* over $\mathbf{Q}[x]$ is in **P**. Neither our probabilistic parallel algorithm nor our deterministic sequential algorithm for the Smith normal form is based on repeated computations of Hermite normal forms as is Kannan and Bachem's algorithm. Our key idea in the parallel algorithm is that though each entry in the Smith normal form is a quotient of two GCDs of possibly exponentially many minors we can quickly produce random linear combinations of these minors whose GCD is with high probability equal to the needed GCD. Unlike our parallel Hermite normal form algorithm our parallel solution for the Smith normal form also provides a practical algorithm superior to previously known methods.

We wish to add two remarks. One can use *HERMITE FORM* over $\mathbf{Q}[x]$ as a tool to solve linear systems over $\mathbf{Q}[x]$ in polynomial-time. Also, however, the fact that solving linear systems over $F[x_1,\ldots,x_v]$, v fixed, is $\mathbf{NC}^1$ reducible to singular linear systems over F is a consequence of Hermann's [8] degree estimates of Hilbert's [9] reduction. See also the appendix of Mayr and Meyer [15] for several corrections to Hermann's proof. Secondly, we cannot hope to provide fast parallel algorithms for *HERMITE FORM* over **Z** and *SMITH FORM* over **Z** unless progress is made on computing GCDs of integers in parallel, a problem easily shown to be $\mathbf{NC}^1$ reducible to 2 by 2 Hermite or Smith normal forms over **Z**.

In this paper we will restrict ourselves to non-singular square input matrices but we note that there are no great difficulties to generalize our approach to rectangular inputs of non-maximal rank (cf. [23]).

2. Parallel Hermite Normal Form Computation

In this section we construct an $\mathbf{NC}^1$-reduction from *HERMITE FORM* over $F[x]$, F a field, to singular linear systems over F. But first we present the necessary definitions and lemmas.

A non-singular n by n matrix H over $F[x]$ is in *Hermite normal form* if it is lower triangular, the diagonal entries are monic, and the entries before the diagonal entry in each row are of lower degree than the diagonal entry. It is well-known that for every non-singular matrix A there exists a unique unimodular matrix U and matrix H in Hermite normal form such that $AU = H$. H is referred to as the Hermite normal form of A. It is fairly clear that Hermite [7] knew the uniqueness though he did not offer a proof. In any case, we need the uniqueness in a stronger form than is usually presented, which we will include as lemma 2.1.

For a matrix A over $F[x]$ let $a_{i,j,k}$ denote the coefficient of x^k in the i, jth entry.

Lemma 2.1: Given the n by n nonsingular matrix A over $F[x]$ with entry degrees less than d, and the vector $(d_1,\ldots,d_n)$ of nonnegative integers, consider the system $AP = G$, where G is lower triangular, and more specifically,

> $p_{i,j}$ are are polynomials of degree less than $nd + \max_{1\le i\le n} d_i$, whose coefficients are unknowns.
> $g_{i,i}$ are monic of degree d_i with lower order coefficients unknowns, and
> for $i > j$, $g_{i,j}$ are polynomials of degree less than d_i with unknowns as coefficients.

This is a system of linear equations over F in the unknown $p_{i,j,k}$ and $g_{i,j,k}$ for which the following statements hold.

1. The system has at least one solution, if and only if each d_i is no less than the degree of the ith diagonal entry of a Hermite normal form of A.

2. If each d_i is exactly the degree of the ith diagonal entry of a Hermite normal form of A, then the system has a unique solution, hence G is the unique Hermite normal form of A and P is unimodular.

Proof: Let H be a Hermite Normal Form of A and U a unimodular matrix such that $AU = H$.

Suppose G and P solve the system for for a given degree vector $(d_1,\ldots,d_n)$. Since U is invertible in $F[x]$, we have $G = AP = HU^{-1}P$. Because G and H are triangular and nonsingular, $U^{-1}P$ must be also. It follows that the degrees d_i must be no less than the degrees of $h_{i,i}$, which proves 1. in one direction.

On the other hand, if for each i, we have $d_i \ge \deg(h_{i,i})$, let $D = diag(x^{d_1-\deg(h_{1,1})}, \ldots, x^{d_n-\deg(h_{n,n})})$. Then the system is solved with $P = UD$ and $G = HD$. Thus 1. is proved if we can show that this solution is expressible within the degree bound given for P. Since $\det(A)P = \mathrm{adj}(A)G$, the degrees in P are bounded by the degrees in $\mathrm{adj}(A)G$, which are bounded by

$(n-1)d + \max_{1 \le i \le n} d_i$.

It remains to show the solution is unique (i.e., $G = H, P = U$) when $d_i = \deg(h_{i,i})$. Let R denote the lower triangular matrix, $U^{-1}P$. It suffices now to show that if G and H are in Hermite normal form and R is a unimodular lower triangular matrix such that $G = HR$, then $R = I$ (and $G = H$). This we do by induction on n, the size of the matrices. Partition this system so that the upper left block is 1 by 1:

$$\begin{bmatrix} g & 0 \\ g^c & G' \end{bmatrix} = \begin{bmatrix} h & 0 \\ h^c & H' \end{bmatrix} \begin{bmatrix} r & 0 \\ r^c & R' \end{bmatrix}$$

We see that $g = hr$, $g^c = h^c r + H' r^c$, and $G' = H'R'$. Now G' and H' are in Hermite normal form, R' is unimodular, so by induction, R' is the $n-1$ by $n-1$ identity matrix and $G' = H'$. Also, since g and h are of the same degree and monic, we have $r = 1$ and $g = h$. If any entry in the column vector r^c is non-zero, let i be the index of the first non-zero entry. Then

$$g_i^c = h_i^c + h'_{i,i} r_i^c. \qquad (\dagger)$$

Since $\deg(h_i^c) < \deg(h'_{i,i}) = d_i$, the degree of the right hand side of (†) is no less than d_i. On the other hand, since $\deg(g_i^c) < \deg(g'_{i,i}) = d_i$, the degree of the left hand side is strictly less, a contradiction. Hence all entries of r^c are 0, and $g^c = h^c$, which completes the proof. □

We now define the size of a matrix A over $F[x]$. Let A be an n by n matrix of d degree polynomials with coefficients in F representable in l bits. Then $\text{size}(A) = n^2 d\, l$, which is the number of bits required to write down A in binary.

Lemma 2.2: For $d_i \le nd$ the linear system of lemma 2.1 consists of $O(n^3 d)$ equations in $O(n^3 d)$ unknowns. Its entries are of size l (0's, 1's, and coefficients of A). □

Now let *LINEAR SYSTEMS* over F be the problem of computing one solution to the (possibly) singular linear System $Ax = b$ or indicating that a solution does not exist, given an n by n matrix A and length n column vector of l bit entries from F. Following Cook [4], we say problem X is $\mathbf{NC}^1$ reducible to problem Y, if there is a uniform family of Boolean circuits for solving X which use oracle circuits to solve Y. For the purpose of defining the depth of such circuits an oracle contributes a depth of $\log(r)$, where r is the fan-in to the oracle. The main theorem of this section now follows.

Theorem 2.1: *HERMITE FORM* over $F[x]$ is $\mathbf{NC}^1$ reducible to *LINEAR SYSTEMS* over F.

Proof: We construct our circuit as follows from processing units at three levels.

1. Let $e = nd \ge \deg(\det(A))$. The input matrix A is passed to each of $n(e+1)$ processors which work in parallel. They are numbered by pairs (i, j) where $1 \le i \le n$ and $0 \le j \le e$. The (i, j) processor constructs from A the appropriate input for a *LINEAR SYSTEM* circuit over F which determines if the system as described in lemma 2.1 can be solved when the degree vector is given

by $d_i = j$ and $d_k = e$, for $k \neq i$. If the oracle produces a solution then *true* is passed to the next step. If the oracle indicates no solution exists, then *false* is passed on. By lemma 2.1 the (i, j) circuit answers *true* just in case the ith diagonal entry of the Hermite normal form has degree less than or equal to j. The depth of the circuit at this point is $O(\log(\text{size}(A)))$, by lemma 2.2.

2. The n circuits numbered 1 through n work in parallel. The ith processor gets input from the $e+1$ circuits of step 1 numbered $(i, 0)$ to (i, e). Its output, d_i, is the minimum j such that the output of processor (i, j) is *true*. Clearly, these circuits have $O(\log(\text{size}(A)))$ depth and polynomial size.

3. One processor recieves the d_i's which are the exact degrees of the diagonal entries of the Hermite normal form. It feeds a *LINEAR SYSTEMS* oracle the system described in lemma 2.1, and by part 3, obtains the desired Hermite normal form. □

Corollary: *HERMITE FORM* over $\mathbf{Q}[x]$ and over $\text{GF}(p)[x]$ is in $\mathbf{NC}^2$.

Proof: The corollary follows from the fact that *LINEAR SYSTEMS* over $\mathbf{Q}$ or $\text{GF}(p)$ is in $\mathbf{NC}^2$ [2], [3], [10] [16]. □

3. Parallel Probabilistic Smith Normal Form Computation

A polynomial matrix S is in *Smith normal form* if it is diagonal, each diagonal entry is monic, and each diagonal entry except the last is a divisor of the succeeding one. If S is equivalent to A, i.e. $A = PSQ$, where P and Q are unimodular, then S is called the Smith normal form of A.

Lemma 3.1: Let A be an n by n non-singular matrix over $F[x]$.

1. There is an n by n matrix S in Smith normal form and unimodular matrices P and Q such that $A = PSQ$.

2. Let s_i^* denote the greatest common divisor of all i by i minors of A. Then the diagonal entries in the Smith normal form of A are $s_{1,1} = s_1^*$, and $s_{i,i} = s_i^*/s_{i-1}^*$, for $i > 1$.

3. Two n by n matrices A and B have the same Smith normal form if and only if they are equivalent.

For a proof see Gohberg, Lancaster, and Rodman [5] or Newman [17]. □

Let C_i^n denote all i element subsets of $\{1, \ldots, n\}$ and let $A_{I,J}$, for $I, J \in C_i^n$, denote the minor of A restricted to the rows in I and columns in J. By the above theorem we could compute the Smith normal form of A by computing $s_i^* = \text{GCD}_{I,J \in C_i^n} A_{I,J}$. The problem is that there are exponentially many i by i minors. To overcome this problem we compute two random linear combinations of $A_{I,J}$ whose GCD is likely to be the wanted GCD. These are the principal i by i minors of two randomly selected matrices equivalent to A. The following lemma shows this suffices. Let $1 \ldots i$ denote the set $\{1, \ldots, i\}$.

Lemma 3.2: let A be an n by n matrix over $F[x]$, and let s_i^* be as in lemma 3.1.2. Let $\mathbf{F}$ be the extension of $F[x]$ by $4n^2$ indeterminants, $\mathbf{F} = F[x][\kappa_{j,k}, \lambda_{j,k}, \mu_{j,k}, \nu_{j,k}]$ Then there exists a polynomial $\pi_i \in \mathbf{F}$ of total degree no more than $4i^2d$ with the following property. For any n by n matrices R, T, U, V over F, $\pi_i(r_{j,k}, t_{j,k}, u_{j,k}, v_{j,k}) \neq 0$ implies that $\mathrm{GCD}(B_{1...i,1...i}, C_{1...i,1...i}) = s_i^*$ where $B = RAT$, $C = UAV$.

Proof: First let the matrices have indeterminate entries, $\mathbf{R} = (\kappa_{j,k})$, $\mathbf{T} = (\lambda_{j,k})$, $\mathbf{U} = (\mu_{j,k})$ and $\mathbf{V} = (\nu_{j,k})$. In this case, we first show $\mathbf{G} = \mathrm{GCD}(\mathbf{B}_{1...i,1...i}, \mathbf{C}_{1...i,1...i}) = s_i^*$ in $\mathbf{F}[x]$ where $\mathbf{B} = \mathbf{R}A\mathbf{T}$ and $\mathbf{C} = \mathbf{U}A\mathbf{V}$, and $\mathbf{F}$ is F with the indeterminates in $\mathbf{R}$, $\mathbf{T}$, $\mathbf{U}$, and $\mathbf{V}$ adjoined. We observe that s_i^* is the only factor of $\mathbf{B}_{1...i,1...i}$ or $\mathbf{C}_{1...i,1...i}$ which lies in $F[x]$. By the Binet-Cauchy formula,

$$\mathbf{B}_{1...i,1...i} = \sum_{K,L \in C_i^n} \mathbf{R}_{1...i,K} A_{K,L} \mathbf{T}_{L,1...i}$$

and

$$\mathbf{C}_{1...i,1...i} = \sum_{K,L \in C_i^n} \mathbf{U}_{1...i,K} A_{K,L} \mathbf{V}_{L,1...i}.$$

Now, clearly the factor of $\mathbf{B}_{1...i,1...i}$ (or $\mathbf{C}_{1...i,1...i}$) in $F[x]$ must divide each $A_{K,L}$. On the other hand, $\mathbf{B}_{1...i,1...i}$ and $\mathbf{C}_{1...i,1...i}$ have no factor in common in $\mathbf{F}[x] \setminus F[x]$ since each involves a different set of indeterminates. This shows our claim on $\mathbf{G}$.
We now consider

$$\mathbf{B}^* = \frac{\mathbf{B}_{1...i,1...i}}{s_i^*} \text{ and } \mathbf{C}^* = \frac{\mathbf{C}_{1...i,1...i}}{s_i^*}.$$

$\mathbf{B}^*$ and $\mathbf{C}^*$ are relatively prime in $\mathbf{F}[x]$, thus $\pi_i = \mathrm{resultant}_x(\mathbf{B}^*, \mathbf{C}^*)$ is non-zero. If $\pi_i(r_{j,k}, t_{j,k}, u_{j,k}, v_{j,k}) \neq 0$ then the polynomials $\mathbf{B}^*(r_{j,k}, t_{j,k}, u_{j,k}, v_{j,k})$ and $\mathbf{C}^*(r_{j,k}, t_{j,k}, u_{j,k}, v_{j,k})$ in $F[x]$ remain relatively prime. (For the theory of resultants, consult for example [21, section 5.8].) Therefore $\mathrm{GCD}(B_{1...i,1...i}, C_{1...i,1...i}) = s_i^*$.
It remains to estimate the degree of π_i. Clearly, $\deg_x(\mathbf{B}^*)$, $\deg_x(\mathbf{C}^*) \leq id$. Their degrees in the other indeterminants are bounded by $2i$, thus the degree of $\pi_i \leq id \times 2i + id \times 2i = 4i^2d$. □

Lemma 3.3: With the notation of the previous lemma, if we select the entries in R, T, U, V randomly from a set $S \subset F$ then the probability

$$\mathrm{Prob}(s_i^* = \mathrm{GCD}(B_{1...i,1...i}, C_{1...i,1...i}), \text{ for all } i,\ 1 \leq i \leq n) \geq 1 - \frac{4n^3d}{\mathrm{cardinality}(S)}.$$

Proof: Let $\pi = \prod_{i=1}^n \pi_i$. We are unlucky only if the randomly selected $r_{j,k}$, $t_{j,k}$, $u_{j,k}$ and $v_{j,k}$ are a zero of π. By a result of Schwartz [19] this happens with probability no more than $\deg(\pi)$/cardinality(S). The degree estimate for π_i now immediately implies that $\deg(\pi) \leq 4n^3d$. □

We now can prove the following theorem.

Theorem 3.1: There is a uniform family of probabilistic circuits of depth $O(\log^2(\text{size}(A)/\varepsilon))$ and polynomial size which compute the Smith normal form over $F[x]$ correctly with probability $1-\varepsilon$. These circuits make $O(n^2 \log(nd/\varepsilon))$ random bit choices. In short, *SMITH FORM* over $\mathbf{Q}[x]$ or GF(p)[x] is in $\mathbf{RNC}^2$.

Proof: By lemma 3.3 the problem reduces to matrix multiplications, determinant and GCD computations. These are in $\mathbf{NC}^2$ [2]. We must make our $4n^2$ random choices from a subset S of $\mathbf{Q}$ for which $4n^3 d/\text{cardinality}(S) < \varepsilon$. The integers less in absolute value than $4n^3 d/\varepsilon$ will do. These are $O(\log(nd/\varepsilon))$ bit numbers.
If the field is too small to allow choice of a sufficiently large set S, S may be chosen from an extension field. Like GCD's, the Smith normal form is an entirely rational form and thus is unchanged if one computes over an extension of the given field. □

Lemma 3.2 remains true if we replace U by an upper triangular and V by a lower triangular matrix, as well as if we do not randomize B. This saves in both matrix multiplications and number of random bits required.

4. Sequential Deterministic Smith Normal Form Computation

The purpose of this section is to establish that *SMITH FORM* over $\mathbf{Q}[x]$ is in $\mathbf{P}$. First we note that it is a consequence of Kannan [12] that *SMITH FORM* over GF(p)[x] is in $\mathbf{P}$, a result on which we will have to depend. We can assume without loss of generality that our input matrix A has integer coefficients. The following lemma is the key to our argument.

Lemma 4.1: Let A be a non-singular n by n matrix over $\mathbf{Q}[x]$ with integer coefficients, $d = \max\{\deg(a_{i,j}) \mid 1 \le i, j \le n\}$, $L = \max\{|a_{i,j,k}| \mid 1 \le i, j \le n,\ 0 \le k \le \deg(a_{i,j})\}$, l_A be the leading coefficient of det(A), and let S be the Smith normal form of A, $d_i = \deg(s_{i,i})$. Then for any prime p which does not divide l_A, exactly one of the following two conditions can occur for $\bar{S}$, the Smith normal form of A mod p.

1. $S \bmod p = \bar{S}$ or
2. $(d_1, \ldots, d_n) \ne (\bar{d}_1, \ldots, \bar{d}_n)$ with $\bar{d}_i = \deg(\bar{s}_{i,i})$.

Furthermore, there exists an integer $B_A \le (n(d+1)L)^{3n^3 d}$ such that if p does not divide B_A condition 1 must occur.

Proof: Let $\bar{s}_i^* = \text{GCD}_{J,K \in C_i^n}(A_{J,K} \bmod p)$, $1 \le i \le n$, $\bar{s}_0^* = 1$. Then by lemma 3.1, $\bar{s}_{i,i} = \bar{s}_i^*/\bar{s}_{i-1}^*$ for $1 \le i \le n$. It is clear that $s_i^* \bmod p$ divides $\bar{s}_i^*$. Let $\bar{e}_i = \deg(\bar{s}_i^*)$, $e_i = \deg(s_i^*)$. Then $\bar{e}_i \ge e_i$, $\bar{e}_0 = e_0 = 0$, $\bar{d}_i = \bar{e}_i - \bar{e}_{i-1}$, $d_i = e_i - e_{i-1}$. Either $\bar{e}_i = e_i$ for all $1 \le i \le n$ or there is a first i such that $\bar{e}_i > e_i$. In the first case, since s_i^* and $\bar{s}_i^*$ are both monic, we have that $s_i^* \bmod p = \bar{s}_i^*$ and hence $s_{i,i} \bmod p = \bar{s}_{i,i}$, In the later case, we have $\bar{d}_i > d_i$.

It remains to establish a condition under which

$$\left(\mathrm{GCD}_{J,K \in C_i^n}(A_{J,K})\right) \bmod p = \mathrm{GCD}_{J,K \in C_i^n}(A_{J,K} \bmod p) \qquad (\dagger)$$

for all $1 \le i \le n$. First we note that for $A_{J,K} = \sum b_j x^j$, $b_j \in \mathbf{Z}$ and $|b_j| \le B = (\sqrt{n}\,(d+1)\,L)^n$ (cf. [6], problem 73-17). Secondly we appeal to the following (cf. [11], lemma 4).

Proposition: If $f_1, \ldots, f_t \in \mathbf{Q}[x]$ are polynomials with integer coefficients and $\deg(f_j) \le e$, then there exists an $\bar{e}$ by $\bar{e}$ determinant $\Delta \in \mathbf{Z} \setminus \{0\}$, $\bar{e} \le 2e$, whose entries are coefficients of the f_j such that for any prime p which does not divide Δ

$$\mathrm{GCD}_{1 \le j \le t}(f_j \bmod p) = (\mathrm{GCD}_{1 \le j \le t}(f_j)) \bmod p$$

Proof: Let $d(x) = \mathrm{GCD}(f_j)$. For any prime p, it is clear that $d \bmod p$ divides $\mathrm{GCD}(f_j \bmod p)$, since $d \bmod p$ divides each $f_j \bmod p$. We show the converse holds for most primes. There exist $s_1, \ldots, s_t \in \mathbf{Q}[x]$ with $\deg(s_j) < e$ such that $\mathrm{GCD}(f_j) = \sum f_j s_j$. Since each term has degree at most $e + (e-1)$, this equation may be viewed as a linear system, $d = Fs$ of at most $2e$ equations over $\mathbf{Q}$ in te varibles, the coefficients of the s_j. The entries of the matrix F are the coefficients of the f_j. Such a linear system has a solution just in case the rank of F is the same as the rank of the augmented matrix (F, d). Since the system has a solution over $\mathbf{Q}$, the rank condition holds. If the rank of $F \bmod p$ is $\bar{e} \le 2e$, then an $\bar{e}$ by $\bar{e}$ minor, Δ, of F must be nonzero. If Δ is nonzero mod p as well, it follows that the rank condition will hold mod p and hence the system will have a solution, s'. Thus $\mathrm{GCD}(f_j \bmod p)$ divides $\sum(f_j \bmod p)(s_j')=d \bmod p$, for polynomials s_j' appropriately constructed from s'. □

Continuing the proof of lemma 4.1, we apply this proposition to $A_{J,K}$ and obtain as the asserted determinant an integer B_i,

$$B_i \le \sqrt{2di}\; B^{2di} \le \sqrt{2dn}\left(n\,(d+1)^2\,L^2\right)^{dn^2} < (n\,(d+1)\,L)^{3dn^2},$$

such that if p does not divide B_i (†) is satisfied for i. It remains to set $B_A = \prod_{i=1}^{n} B_i$. □

The deterministic algorithm is now easy to describe. First we select $k = 2\left\lceil \log_2(l_A(n\,(d+1)\,L)^{3n^3d}) \right\rceil \ge 2\lceil \log_2(l_A B_A) \rceil$ primes p_j and compute for all primes not dividing l_A the Smith normal form $\bar{S}_j$ of $A \bmod p_j$. We note that the kth prime p_k is $\le k \log(k)$, $k \ge 6$, which makes this step a polynomial-time process. Also more than half of the primes considered do not divide $l_A B_A$. Hence by the above lemma a majority of the $\bar{S}_j$ must possess the same diagonal-degree vector, say these mod p_j, $j \in J$. Also by the lemma $\bar{S}_j$, $j \in J$, is an image of S. By Chinese remaindering we compute

$$\tilde{S} \equiv S \bmod \tilde{p},\ \tilde{p} = \prod_{j \in J} p_j.$$

It remains to recover the coefficients $s_{i,i,k}$ from their modular images $\tilde{s}_{i,i,k}$. We first observe that

the $s_{i,i}$ are monic factors of det(A) over $\mathbf{Q}[x]$. Therefore by Gauss' lemma the denominators of $s_{i,i,k}$ are factors of l_A and hence relatively prime to $\tilde{p}$. We now claim that

$$s_{i,i,k} = \frac{l_A \tilde{s}_{i,i,j} \bmod \tilde{p}}{l_A},$$

where the modulus in the numerator is taken balanced. The only problem could be that $\tilde{p}$ were too small to capture $l_A \times$ the numerator of $s_{i,i,k}$. But the integral coefficients of factors of det(A) are absolutely bounded by $2^{nd} B$ (see [14], Section 4.6.2, Exercise 20). Now clearly $2 l_A 2^{nd} B < \tilde{p}$ and we have the following theorem.

Theorem 4.1: *SMITH FORM* over $\mathbf{Q}[x]$ is in **P**.

5. Rational Canonical form and Similarity

If A is a matrix over a field F, then the diagonal entries of the Smith normal form of $xI - A$ (over $F[x]$) are the invariant factors of A. The invariant factors characterize A up to similarity and their companion matrices form the diagonal blocks of the rational canonical form R of A. Thus we can compute *RATIONAL FORM* in $\mathbf{RNC}^2$ and in **P**. Furthermore, we can compute the similarity transform U such that $UAU^{-1} = R$, whereas for the Smith normal form S such that $PAQ = S$, we did not obtain P and Q. Knowing U, we can verify that $UAU^{-1} = R$, Thus the probabilistic algorithm for Rational normal form is of Las Vegas type (controllably small probability of no result), whereas the Smith normal form algorithm was of Monte Carlo type (controllably small probability of incorrect result).

To compute the transform U, first compute R via the Smith form of $xI - A$, as indicated above. Then solve the linear system $UA = RU$. An arbitrary U satisfying this equation will not do, as it may be singular. However, we may do the following. We compute a basis U_1, ..., U_k of the solution space. Let λ_1, ..., λ_k be indeterminants and let

$$\pi(\lambda_1, \ldots, \lambda_k) = \det(\sum_{i=1}^{k} \lambda_i U_i)$$

We choose r_1, ..., r_k at random from F and let

$$U = \sum_{i=1}^{k} r_i U_i$$

Then U is nonsingular unless $\pi(r_1, ..., r_k) = 0$. We know that π is not identically zero since if R is the rational form of A, then by definition, a nonsingular U such that $UA = RU$ must exist. By Schwartz' result [19] the probability that we unluckily obtain a singular U is less than $\deg(\pi)/s$, where s is the size of the set from which we choose the components of (r_1, ..., r_k). Thus, if det(U) is nonzero R is a verified rational canonical form of A. If it is zero then we return no solution. Either we were unlucky in computing R via the probabilistic Smith normal form algorithm, or we were unlucky in computing U. If F is finite, and a larger s is desired, the r_i may be chosen from an extension of F.

More details on this and the construction of the Jordan normal form can be found in [22].

6. Conclusion

In the meantime, we have discovered a Las Vegas solution for the Smith normal form problem of polynomial matrices [23]. This solution hinges greatly on the Hermite normal form process, as opposed to the Monte-Carlo solution proposed here. Its analysis, however, is similar to the one here. The new algorithm also finds the multipliers. In the future we will carry out practical experiments with our randomized algorithms.

References

[1] Bachem A., and R. Kannan, "Applications of Polynomial Smith Normal Calculations," Report #78119-OR, Institut fur Ökonometrie und Operations Research Universität Bonn, 1978.

[2] Borodin, A., von zur Gathen, J., and Hopcroft, J. "Fast parallel matrix and GCD computations," Information and Control, Vol. 52, No. 3, March 1982, pp. 241-256.

[3] Borodin, A., S.A. Cook, and N. Pippenger, "Parallel computation for well-endowed rings and space-bounded probablistic machines," Tech. Report 162/83, University of Toronto, April 1983.

[4] Cook, S.A., "The Classification of problems which have fast parallel algorithms," Proc. Int. Conf. Foundations of Computation Theory, Borgholm, 1983, Springer Verlag Lecture Notes in Computer Science, Vol. 158, pp 78-93.

[5] Gohberg, I., P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press, 1982.

[6] Goldstein, A.J., and R.L. Graham, "A Hadamard-Type Bound on the coefficients of a determinant of polynomials," SIAM Review, Vol. 16, 1974, pp. 394-395.

[7] Hermite, C., "Sur l'introduction des variables continues dans la theorie des nombres," J. reine angew. Math., Vol. 41, 1851, pp. 191-216.

[8] Hermann, G., "Die Frage der endlich vielen Schritte in der Theorie der Polynomideale," Math. Ann. 95, 1922, pp.736-788.

[9] Hilbert, D., "Über die Theorie der algebraischen Formen," Math. Ann., Vol. 36, 1890, pp. 473-534.

[10] Ibarra, O.H. , S. Moran, and L.E. Rosier, "A note on the parallel complexity of computing the rank of order *n* matrices," Vol. 11, Information Processing Letters, December 1980, pp. 162.

[11] Kaltofen, E., "Effective Hilbert Irreducibility," EUROSAM 1984, Lecture Notes in Comp. Sci., Vol. 174, Springer, 1984, pp. 277-284.

[12] Kannan, R., “Polynomial-time algorithms for solving systems of linear equations over polynomials,” Theoretical Computer Science, Vol. 39, 1985, pp. 69-88.

[13] Kannan, R., and A. Bachem, “Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix,” SIAM J. Computing, Vol. 8, 1981, pp. 499-507.

[14] Knuth, D.E., *The Art of Programming*, Vol. 2, *Seminumerical Algorithms*, 2nd edition, Addison Wesley, 1981.

[15] Mayr, E.W. and A.R. Meyer, “The Complexity of the Word Problems for Commutative Semigroups and Polynomial Ideals,” Advances in Mathematics, Vol. 46, December 1982, pp. 305-329.

[16] Mulmuley, K., “A Fast Parallel Algorithm to Compute the Rank of a Matrix over an Arbitrary Field,” Manuscript, 1985.

[17] Newman, M., *Integral Matrices*, Academic Press, 1972.

[18] Ramachandran, V., “Exact Reduction of a Polynomial Matrix to the Smith Normal Form,” IEEE Transanctions on Automatic Control, AC-24 (1979), 638-641.

[19] Schwartz, J.T., “Fast probabilistic algorithms for verification of polynomial identities,” J. ACM, 27 (1980), 701-717.

[20] Smith, H.J.S., “On systems of linear indeterminate equations and congruences,” Philosophical Transactions, 151(1861), pp. 293-326.

[21] Van der Waerden, B.L., *Algebra*, Frederic Ungar Publ., 1970

[22] Kaltofen, E., M. Krishnamoorthy, and B.D. Saunders, “Fast Parallel Algorithms for Similarity of Matrices,” Proc. 1986 ACM Symp. Symbolic Algebraic Comput., pp. 65-70.

[23] Kaltofen, E., M. Krishnamoorthy, and B.D. Saunders, “Mr. Smith Goes to Las Vegas, Randomized Parallel Computation of the Smith Normal Form of Polynomial Matrices,” Manuscript, June 1987.