An Algorithm to Calculate the Kernel of Certain Polynomial Ring Homomorphisms

Fausto Di Biase and Rüdiger Urbanke

CONTENTS

We propose an improvement upon the standard algorithm for computing the kernel of a polynomial map, assuming that the map sends monomials into monomials. Rather than computing a Gröbner basis in the joint polynomial ring, and then selecting only the elements of interest, we show that a moderate number of iterations of the Buchberger algorithm in the variables of the domain suffices.

1. INTRODUCTION

We are interested in calculating a finite basis for the kernel of a ring homomorphism $\pi : K_x \to K_y$ between polynomial rings $K_x := K[x_1, \ldots, x_n]$ and $K_y := K[y_1, \ldots, y_m]$. It is well known [Adams and Loustaunau 1994; Conti and Traverso 1991] that this can be accomplished applying the Buchberger algorithm [Buchberger 1985] over the polynomial ring $K_{x,y} := K[x_1, \ldots, x_n, y_1, \ldots, y_m]$.

Unfortunately, the complexity of the Buchberger algorithm is a strongly increasing function of the number of variables. Hence, it would be useful to find an algorithm operating on K_x alone. The main result of this paper is that this can indeed be done in the special case in which the map π is the extension of a semigroup homomorphism. We will see that in this case a moderate number (bounded by $\lfloor \frac{1}{2} n \rfloor$) of Buchberger algorithms over K_x is sufficient to find a basis for $\ker \pi$, and, hence, that for a large number of variables the proposed algorithm will be more efficient than the standard algorithm.

For general information on Gröbner bases, see [Buchberger 1985; Cox et al. 1991].

The proposed algorithm has many applications, e.g., in the area of integer programming [Conti

This work was carried out while the authors were at Washington University, St. Louis, in the Mathematics and Electrical Engineering Departments, respectively.

Di Biase thanks the C.N.R. (Italy) for support, through grants 203.01.55 and 203.01.63.

1058-6458/96 $0.50 per page

and Traverso 1991; Natraj et al.; Thomas; Hosten and Sturmfels 1994] as well in the realm of sampling from conditional distributions [Diaconis and Sturmfels].

We start by shortly reviewing the standard solution in Section 2. In Section 3 we then present the new algorithm together with some examples. Finally, in Section 4 the running times of these two algorithms are compared and the results are discussed.

2. THE SOLUTION BY MEANS OF A GRÖBNER BASIS OVER $K_{x,y}$

Let $f_j = \pi(x_j) \in K_y$ and $\gamma_j = x_j - f_j \in K_{x,y}$, for $j = 1, \dots, n$. Denote by $\langle\{\dots\}\rangle$ the ideal generated by $\{\dots\}$.

Lemma 2.1. *Let G be a (reduced) Gröbner basis for $\langle\{\gamma_1, \dots, \gamma_n\}\rangle$ in $K_{x,y}$ with respect to a term order that eliminates the y variables. Then $G' = G \cap K_x$ is a (reduced) Gröbner basis for* $\ker\pi$.

Proof. The Elimination Theorem [Cox et al. 1991, p. 114] implies that G' is a (reduced) Gröbner basis for $\langle\{\gamma_1, \dots, \gamma_n\}\rangle \cap K_x$. The claim then follows since $\langle\{\gamma_1, \dots, \gamma_n\}\rangle$ is equal to the kernel of the (unique) homomorphic extension $\pi_e : K_{x,y} \to K_y$ of π, for which $\pi_e(x_j) = f_j$ and $\pi_e(y_i) = y_i$. □

Example 2.2. Let $\pi : K[x_1, \dots, x_4] \to K[y_1, y_2]$ map $x_1, \dots, x_4$ into y_1^3, $y_1^2 y_2$, $y_1 y_2^2$, and y_2^3, respectively, so that $\gamma_1 = x_1 - y_1^3$, etc. Then

$$\begin{aligned} G = \{&-x_3^2 + x_2x_4, -x_2x_3 + x_1x_4, -x_2^2 + x_1x_3,\\ &x_4 - y_2^3, -x_4y_1 + x_3y_2, -x_3y_1 + x_2y_2,\\ &-x_2y_1 + x_1y_2, x_3 - y_1y_2^2, x_2 - y_1^2y_2, x_1 - y_1^3\} \end{aligned}$$

is the reduced Gröbner basis for $\langle\{\gamma_1, \dots, \gamma_4\}\rangle$ with respect to lex order, $y_1 > y_2 > x_1 > \dots > x_4$. Hence,

$$G' = \{-x_3^2 + x_2x_4,\ -x_2x_3 + x_1x_4,\ -x_2^2 + x_1x_3\}$$

is the reduced Gröbner basis for $\ker\pi$ with respect to lex order, $x_1 > \dots > x_4$.

Example 2.3. Let $\pi : K[x_1, \dots, x_6] \to K[y_1, \dots, y_4]$ map $x_1, \dots, x_6$ to, respectively, $y_1^2y_2^3y_3^5$, $y_1^5y_2^2y_3^3y_4$, $y_1^4y_2^4y_3^2y_4$, $y_1y_3^2y_4^4$, $y_1y_3^5y_4^3$, and $y_2^2y_4^2$. The reduced Gröbner basis G for $\langle\{\gamma_1, \dots, \gamma_6\}\rangle$ with respect to lex order $y_1 > \cdots y_4 > x_1 > \cdots > x_6$ consists of 1180 elements, and

$$\begin{aligned} G' = \{&x_3^{10}x_4^6 - x_2^9x_5x_6^{11}, x_1^{14}x_4^{20} - x_2x_3^6x_5^{19}x_6^8,\\ &x_3^{16}x_5^{18} - x_1^{14}x_2^8x_4^{14}x_6^3, x_3^{26}x_5^{17} - x_1^{14}x_2^{17}x_4^8x_6^{14},\\ &x_3^{36}x_5^{16} - x_1^{14}x_2^{26}x_4^2x_6^{25}, x_3^{46}x_4^4x_5^{15} - x_1^{14}x_2^{35}x_6^{36}\} \end{aligned}$$

is the reduced Gröbner basis for $\ker\pi$ with respect to lex order $x_1 > \cdots > x_6$.

Note that the above procedure requires the calculation of a Gröbner basis G of $\langle\{\gamma_i\}\rangle$ in $K_{x,y}$, but the actual solution G' is then just a small subset of G, namely $G' = G \cap K_x$. This is especially apparent in the second example, where $|G| = 1180$ and $|G'| = 6$.

3. THE SOLUTION BY MEANS OF GRÖBNER BASES OVER K_x

Here we present an alternative way to calculate $\ker\pi$ when $\pi(x_j)$ is a monomial in K_y for every $j = 1, \dots, n$. The solution is gradually built up by a repeated application of the Buchberger algorithm over K_x (as opposed to $K_{x,y}$). The efficiency of the proposed algorithm is based on the (empirical) fact that the complexity of the Buchberger algorithm grows strongly in the number of variables, so that for a large number of variables it is more efficient to calculate a moderate number of Gröbner bases over K_x instead of one over $K_{x,y}$. This is especially true for the memory requirements of the proposed algorithm (see Example 2.3).

Let $\mathcal{M}_x$ denote the set of all monomials in K_x, and likewise $\mathcal{M}_y$. Let $f_j = \pi(x_j)$ and assume that $f_j \in \mathcal{M}_y$, for $j = 1, \dots, n$. Note that all the information about π is contained in the $m \times n$ matrix M, with nonnegative integer entries, given by the exponents of the monomials $f_1, \dots, f_n$.

We use the usual compact multi-index notation for monomials: e.g., for $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, the symbol x^α denotes the monomial ${x_1}^{\alpha_1} \dots {x_n}^{\alpha_n}$. Denote by log the isomorphism between $\mathcal{M}_x$ and

$\mathbb{N}^n$ given by $\log x^\alpha := \alpha$ for $\alpha \in \mathbb{N}^n$. There is a similar map from $\mathcal{M}_y$ to $\mathbb{N}^m$, which, by a slight abuse of notation, we will denote by log also.

Let $M \in \mathbb{N}^{m\times n}$ be defined by $M_{i,j} = (\log f_j)_i$, where $M_{i,j}$ is the entry of M in row i, column j. This matrix defines a $\mathbb{Z}$-linear mapping $\pi_* : \mathbb{Z}^n \to \mathbb{Z}^m$, by the usual row-by-column multiplication:

$$\pi_*(u) := Mu.$$

A $\mathbb{Z}$-basis for $\ker \pi_*$ can be calculated using a variant of an algorithm that calculates the Smith normal form of an integer basis [Cohen 1993, p. 72]. The connection between $\ker \pi$ and $\ker \pi_*$ is conveyed by a map $\varphi : \mathbb{Z}^n \to K_x$, which will be presently defined.

For $u \in \mathbb{Z}$ define u^+ and u^- by $u^+ = \max(u, 0)$ and $u^- = \max(-u, 0)$. These definitions extend naturally to elements in $\mathbb{Z}^n$ if we apply them componentwise. Note that if $u \in \mathbb{Z}^n$ then $u^+, u^- \in \mathbb{N}^n$, and that $(-u)^+ = u^-$ and $(-u)^- = u^+$. Now for any $u \in \mathbb{Z}^n$ we define

$$\varphi(u) := x^{u^+} - x^{u^-}.$$

Any binomial (difference of monomials) $p = x^\alpha - x^\beta$, where $\alpha, \beta \in \mathbb{N}^n$, can be written (in a unique way) as $p = x^\alpha - x^\beta = m_p\varphi(u_p)$, where $m_p \in \mathcal{M}_x$ and $u_p \in \mathbb{Z}^n$. We will write p^+ for x^α, and p^- for x^β. Then clearly $p^+ = x^\alpha = m_p x^{u_p^+}$ and $p^- = x^\beta = m_p x^{u_p^-}$. Let $\le'$ denote the *natural partial order* on $\mathbb{N}^n$ obtained by forming the product of n copies of $\mathbb{N}$ with its natural order. If $\alpha, \beta \in \mathbb{N}^n$, we denote by $\alpha \vee \beta$ the $\le'$-smallest element of $\mathbb{N}^n$ such that $\alpha \le' \alpha \vee \beta$ and $\beta \le' \alpha \vee \beta$. In symbols,

$$(\alpha \vee \beta)_j = \max\{\alpha_j, \beta_j\} \qquad \text{for } j = 1, \dots, n.$$

We set $x^\alpha \vee x^\beta = x^{\alpha\vee\beta}$.

Lemma 3.1. *Let p and q be binomials in K_x. Then*

$$\frac{p^- \vee q^+}{p^-} p + \frac{p^- \vee q^+}{q^+} q = m_{p,q}\, \varphi(u_p + u_q) \tag{3.1}$$

for some $m_{p,q} \in \mathcal{M}_x$.

Proof. Note that $u_p - (u_p + u_q)^+ = -u_q - (u_p + u_q)^-$. Therefore the left-hand side equals

$$\begin{aligned}
&\frac{p^- \vee q^+}{p^-}(p^+ - p^-) + \frac{p^- \vee q^+}{q^+}(q^+ - q^-)\\
&\quad= \frac{p^- \vee q^+}{p^-} p^+ - \frac{p^- \vee q^+}{q^+} q^-\\
&\quad= x^{u_p + (u_p^- + \log(m_p)) \vee (u_q^+ + \log(m_q))}\\
&\qquad\qquad - x^{-u_q + (u_p^- + \log(m_p)) \vee (u_q^+ + \log(m_q))}\\
&\quad= x^\alpha \left(x^{(u_p+u_q)^+} - x^{(u_p+u_q)^-}\right)\\
&\quad= x^\alpha \varphi(u_p + u_q),
\end{aligned}$$

where

$$\begin{aligned}
\alpha &= u_p - (u_p + u_q)^+ + (u_p^- + \log(m_p)) \vee (u_q^+ + \log(m_q))\\
&= -u_q - (u_p + u_q)^- + (u_p^- + \log(m_p)) \vee (u_q^+ + \log(m_q)).
\end{aligned}$$

To see that $\alpha \in \mathbb{N}^n$ note that $(u_p^- + \log(m_p)) \vee (u_q^+ + \log(m_q))$ is increasing (with respect to $\le'$) in m_p and m_q. Hence, it suffices to show that $\alpha \in \mathbb{N}^n$ in the case $m_p = m_q = 1$. Using the facts that $u_p^- \vee u_q^+ = u_p^- + (u_q^+ - u_p^-)^+$ and $(u_p + u_q)^+ = (u_p^+ - u_q^-)^+ + (u_q^+ - u_p^-)^+$, we get that $\alpha = u_p^+ - (u_p^+ - u_q^-)^+$ and, hence, $\alpha \in \mathbb{N}^n$. □

Note that the S-polynomial [Cox et al. 1991, p. 82] of p and q, as well as the reduction of p with respect to q (if possible), can be written in the form of equation (3.1). This shows that if we calculate the Gröbner basis of a set of elements of the form $m_i\varphi(u_i)$, where $m_i \in \mathcal{M}_x$ and $u_i \in \mathbb{Z}^n$, each element of this Gröbner basis will also have this form, since this calculation can be done by computing a sequence of S-polynomials (possibly) followed by reductions. This can be seen in Examples 2.2 and 2.3 (each γ_j is a binomial in $K_{x,y}$). We also have the following special case. (The support of $\alpha \in \mathbb{N}^n$, denoted by $\operatorname{supp}\alpha$, is the set of indices for which the corresponding component of α is not zero.)

Corollary 3.2. *Let $p = \varphi(u_p), q = \varphi(u_q)$ for some $u_p, u_q \in \mathbb{Z}^n$. If $\operatorname{supp} u_p^+$ and $\operatorname{supp} u_q^-$ are disjoint (or $\operatorname{supp} u_p^-$ is disjoint from $\operatorname{supp} u_q^+$) then*

$$\varphi(u_p + u_q) \in \langle \varphi(\{u_p, u_q\})\rangle.$$

Proof. From Lemma 3.1 with $m_p = m_q = 1$ we have (using the same notation as in the previous proof) $\alpha = u_p^+ - (u_p^+ - u_q^-)^+$. The hypothesis on the supports implies that $(u_p^+ - u_q^-)^+ = u_p^+$ and, hence, $\alpha = 0$. $\square$

Example 3.3. Let $u = (1, -2, 3, -1) = u^+ - u^- = (1, 0, 3, 0) - (0, 2, 0, 1)$ and $v = (2, 1, 0, -2) = v^+ - v^- = (2, 1, 0, 0) - (0, 0, 0, 2)$. Since

$$\operatorname{supp} u^+ \cap \operatorname{supp} v^- = \{1, 3\} \cap \{4\} = \varnothing,$$

it follows from Corollary 3.2 that

$$\varphi(u + v) = \varphi((3, -1, 3, -3)) = x_1^3 x_3^3 - x_2 x_4^3$$

belongs to $\langle \{x_1 x_3^3 - x_2^2 x_4,\ x_1^2 x_2 - x_4^2\} \rangle$. Indeed, this binomial equals $x_1^2(x_1 x_3^3 - x_2^2 x_4) + x_2 x_4 (x_1^2 x_2 - x_4^2)$.

In the next theorem we see that the connection between $\ker \pi_*$ and $\ker \pi$ rests on the fact that $\ker \pi$ is the ideal generated by $\varphi(\ker \pi_*)$.

Theorem 3.4. $\ker \pi = \langle \varphi(\ker \pi_*) \rangle$.

Proof. It is simply based on a telescopic identity. Details can be found in [Herzog 1970]. $\square$

If K is a basis for $\ker \pi_*$ consisting of k elements (to simplify notation, we will use K to denote a basis for $\ker \pi_*$ as well as the matrix in $\mathbb{Z}^{k \times n}$ whose rows are the vectors in K), then it is not true in general that $\varphi(K)$ will be a set of generators for $\langle \varphi(\operatorname{span} K) \rangle$, as the next example shows.

Example 3.5. Let π be as in Example 2.2. Then

$$M = \begin{pmatrix} \log(f_1)_1 & \log(f_2)_1 & \log(f_3)_1 & \log(f_4)_1 \\ \log(f_1)_2 & \log(f_2)_2 & \log(f_3)_2 & \log(f_4)_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

Calculating a $\mathbb{Z}$-basis for $\ker \pi_*$ we get

$$K = \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 2 & -1 & 0 \end{pmatrix}.$$

Hence, $\varphi(K) = \{x_1 x_4 - x_2 x_3,\ x_2^2 - x_1 x_3\}$ and, calculating the reduced Gröbner basis with respect to lex order, $x_1 > \cdots > x_4$, we get the set

$$G = \{x_2 x_3^2 - x_2^2 x_4,\ -x_2 x_3 + x_1 x_4,\ x_2^2 - x_1 x_3\}.$$

Since this (reduced and therefore unique) Gröbner basis does not equal the Gröbner basis we calculated in Example 2.2, we conclude that $\langle \varphi(K) \rangle$ is strictly contained in $\langle \varphi(\operatorname{span} K) \rangle$.

Example 3.6. There are many equivalent choices for K. More precisely, if $A \in \mathbb{Z}^{k \times k}$ and $\det A = \pm 1$, then $K' = AK$ is an equivalent basis ($\operatorname{span} K' = \operatorname{span} K$), and we write $K' \sim K$. Conversely, any two equivalent bases are related in this way. Hence, we may ask if there always exists a $K' \sim K$ such that $\langle \varphi(K') \rangle = \langle \varphi(\operatorname{span} K) \rangle$. Again the previous example shows that this is not true.

But there is an important special case in which $\varphi(K)$ is already sufficient to generate $\langle \varphi(\operatorname{span} K) \rangle$.

Theorem 3.7. *Let $K \in \mathbb{N}^{k \times n}$. Then*

$$\langle \varphi(K) \rangle = \langle \varphi(\operatorname{span} K) \rangle.$$

Proof. Let g_i, $i = 1, \ldots, k$, denote the rows of K. It suffices to prove that if $\varphi(v) \in \langle \varphi(K) \rangle$ and $i \in \{1, \ldots, k\}$ then $\varphi(v \pm g_i) \in \langle \varphi(K) \rangle$, since any $u \in \operatorname{span} K$ can be achieved in this way by starting with $v = 0$. Hence, assume $\varphi(v) \in \langle \varphi(K) \rangle$. Note that $g_i \in \mathbb{N}^n$, so that $\operatorname{supp} g_i^- \cap \operatorname{supp} v^+ = \varnothing$. Therefore, by Corollary 3.2,

$$\varphi(v + g_i) \in \langle \varphi(\{v, g_i\}) \rangle \subseteq \langle \varphi(K) \rangle.$$

To prove that $\varphi(v - g_i) \in \langle \varphi(K) \rangle$, note that $v - g_i = v + (-g_i)$, and that

$$\operatorname{supp}(-g_i)^+ \cap \operatorname{supp} v^- = \operatorname{supp} g_i^- \cap \operatorname{supp} v^- = \varnothing.$$

$\square$

In general, given a $K \in \mathbb{Z}^{k \times n}$ there will not exist a $K' \sim K$ such that $K' \in \mathbb{N}^{k \times n}$. Nevertheless we can always find an equivalent basis with all base vectors lying in the same orthant, as the next lemma shows.

Lemma 3.8. *Let $K \in \mathbb{Z}^{k \times n}$. Then there exists a $K' \sim K$ such that each column of K' is either in $\mathbb{N}^k$ or in $(-\mathbb{N})^k$.*

Proof. First note that, if the j-th column of K is the zero k-tuple, the j-th column of any $K' \sim K$ is an element of $\mathbb{N}^k$ (being also the zero vector). Hence, without loss of generality, we can assume that each column of K contains at least one nonzero entry. If we can show that there exists a $K' \sim K$ such that some row of K' has all nonzero entries then by adding suitable (positive) multiples of this row to all other rows (thereby not changing the row space) we can generate the desired equivalent matrix. To see that such a K' exists, consider the set

$$\{i : 1 \le i \le n \text{ and } |K_{j,i}| > 0\},$$

called the support of the j-th row of K and denoted by $\operatorname{supp} g_j$. Define K' by $g'_1 = g_1$ and $g'_j = g_j + a_{j-1} g'_{j-1}$ for $j = 2, \dots, k$, where a_{j-1} is any integer such that $\operatorname{supp} g'_j = \operatorname{supp} g_j \cup \operatorname{supp} g'_{j-1}$ (for instance, $a_{j-1} = 1 + \max_i |g_{j,i}|$). Clearly $K' \sim K$ and

$$\operatorname{supp} g'_k = \bigcup_{j=1}^{k} \operatorname{supp} g_j = \{1, \dots, n\}.$$

Therefore all components of g'_k are nonzero. □

Example 3.9. Let π be as in Example 2.2 and Example 3.5. Then

$$\begin{aligned} M &= \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 2 & -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -3 & 0 & 1 \\ 1 & -2 & 1 & 0 \end{pmatrix} \end{aligned}$$

is an equivalent basis for $\ker \pi_*$ with both rows in the same orthant.

Lemma 3.10. *Let $K \in \mathbb{Z}^{k \times n}$, and assume there exists a finite set $U \subset \operatorname{span} K$ such that $\langle \varphi(U) \rangle = \langle \varphi(\operatorname{span} K) \rangle$. If G is the reduced Gröbner basis for $\langle \varphi(U) \rangle$ (with respect to some fixed order $<$) then $G = \varphi(\tilde{U})$ for some $\tilde{U} \subset \operatorname{span} K$.*

Proof. By the remark following Lemma 3.1, each element g in the reduced Gröbner basis has the form $g = m\varphi(u)$ for some $m \in \mathcal{M}_x$ and $u \in \operatorname{span} K$. Assume that $m \neq 1$ for some $g \in G$. Now, $\varphi(u) \in \langle \varphi(U) \rangle$ implies that $\varphi(u)$ (and, hence, g) is reducible by $G \setminus \{g\}$, contradicting the fact that G was already reduced. □

For $j \in \{1, 2, \dots, n\}$ define $T_j : \mathbb{Z}^n \to \mathbb{Z}^n$ as the operator that switches the sign of the jth component of the vectors in $\mathbb{Z}^n$. Further, if $p \in K_x$ has the form $p = \varphi(u)$ for some $u \in \mathbb{Z}^n$, we define $T_j(p) = \varphi(T_j u)$.

Theorem 3.11. *Let $K \in \mathbb{Z}^{k \times n}$ and assume that there exists a finite set $U \subset \operatorname{span} K$ such that $\langle \varphi(U) \rangle = \langle \varphi(\operatorname{span} K) \rangle$. If G is the reduced Gröbner basis for $\langle \varphi(U) \rangle$, with respect to a term order that eliminates x_j, then $\langle T_j G \rangle = \langle \varphi(\operatorname{span}(T_j K)) \rangle$.*

Proof. First observe that, by Lemma 3.10, there is a finite subset $\tilde{U} \subset \operatorname{span} K$ such that $G = \varphi(\tilde{U})$. Thus $T_j G = \varphi(T_j \tilde{U})$ is well defined. Moreover, since $T_j \tilde{U} \subset \operatorname{span}(T_j K)$, we have

$$\langle T_j G \rangle \subset \langle \varphi(\operatorname{span}(T_j K)) \rangle.$$

To prove the other inclusion, let $u \in \operatorname{span} K$. Since G is a Gröbner basis, there exists a $g \in G$ that gives rise to a head reduction of $\varphi(u)$. Without loss of generality, we may assume that $\varphi(u) = x_j^a p - q$ and $g = x_j^b r - s$, where $a, b \in \mathbb{N}$, $a \ge b$ and p, q, r, s are monomials in $K[x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n]$. Thus $p = rm$ for some $m \in K[x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n]$ and

$$\begin{aligned} \varphi(u) &= x_j^a p - q = (x_j^b r - s) x_j^{a-b} m + (x_j^{a-b} s m - q) \\ &= g x_j^{a-b} m + \varphi(\tilde{u}) \tilde{m}, \end{aligned}$$

where $\tilde{m}$ is a monomial in

$$K[x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n],$$

$\tilde{u} \in \operatorname{span} K$ by Lemma 3.1, and the leading term of $\varphi(\tilde{u})$ is strictly smaller (with respect to the chosen term order) than the leading term of $\varphi(u)$. We have $\varphi(T_j u) = p - x_j^a q$ and $T_j g = r - x_j^b s$. Hence,

$$\begin{aligned} \varphi(T_j u) &= p - x_j^a q = (r - x_j^b s) m + (sm - x_j^{a-b} q) x_j^b \\ &= (T_j g) m + \varphi(T_j \tilde{u}) x_j^b \tilde{m}. \end{aligned}$$

We see that $\varphi(T_j u) \in \langle T_j G \rangle$ will follow from

$$\varphi(T_j \tilde{u}) \in \langle T_j G \rangle.$$

The latter follows by induction, from the observation that the same procedure can be applied to

$\varphi(\tilde{u})$ (since $\tilde{u} \in \operatorname{span} K$) and that the leading term of $\varphi(\tilde{u})$ is strictly smaller than the leading term of $\varphi(u)$. $\square$

We are now ready to describe our algorithm to calculate $\ker \pi$. Let K be a basis for $\ker \pi_*$. By Lemma 3.8 there exists an equivalent basis K' such that each column of K' is either in $\mathbb{N}^n$ or in $(-\mathbb{N})^n$. Let $J \subseteq \{1, 2, \dots, n\}$ be the index set of all columns with negative entries, and let K'_J be the matrix obtained from K' by reversing all signs in the columns indexed by J. By Theorem 3.7,

$$\langle \varphi(K'_J) \rangle = \langle \varphi(\operatorname{span} K'_J) \rangle.$$

If $J = \varnothing$ we are done. If $J \neq \varnothing$, let j be any element of J. Theorem 3.11 enables us to derive from $\varphi(K'_J)$ a finite set of generators for $\langle \varphi(\operatorname{span} K'_{J\setminus\{j\}}) \rangle$. Compute the Gröbner basis for $\varphi(K'_J)$ with respect to a term order that eliminates x_j and apply the operator T_j to it. Proceeding recursively, we can calculate a finite set of generators for $\varphi(\operatorname{span} K'_J)$, which by Theorem 3.4 equals $\ker \pi$.

Summary of the Algorithm. 1. Calculate a basis K for $\ker \pi_*$.
2. Find an equivalent basis K' such that all rows of K' lie in the same orthant.
3. Let J be the index set of all columns with negative entries and let K'_J be the matrix obtained from K' by reversing the signs of the columns indexed by J.
4. Let $G_J = \varphi(K'_J)$.
5. Until $J = \varnothing$, repeat this: Take $j \in J$ and let $G_{J\setminus\{j\}}$ be the result of T_j operating on the reduced Gröbner basis for G_J with respect to a term order that eliminates x_j; then let $J \leftarrow J \setminus \{j\}$.
6. Output $G_\varnothing$, a generating set for $\ker \pi$.

Example 3.12. Let π be as in Examples 2.2, 3.5, and 3.9. Then $G_{\{2\}} = \varphi(K'_{\{2\}}) = \varphi(T_2 K') = \{x_1 x_2^2 x_3 - 1,\ x_1^2 x_2^3 x_4 - 1\}$. Calculating a Gröbner basis of $G_{\{2\}}$ with respect to lex order, $x_2 > x_1 > x_3 > x_4$, we get

$$\{-x_4^2 x_1 + x_3^3,\ x_4 - x_2 x_3^2,\ -x_4 x_1 x_2 + x_3,\ -1 + x_1 x_2^2 x_3\}.$$

Hence, $G_\varnothing := T_2 G_{\{2\}} = \{-x_4^2 x_1 + x_3^3,\ x_2 x_4 - x_3^2,\ -x_4 x_1 + x_2 x_3,\ -x_2^2 + x_1 x_3\}$ is a generating set for $\ker \varphi$. Calculating a reduced Gröbner basis of $G_\varnothing$ with respect to lex order $x_1 > \cdots > x_4$, we obtain $\{-x_3^2 + x_2 x_4,\ -x_2 x_3 + x_1 x_4,\ -x_2^2 + x_1 x_3\}$, which agrees with our previous result derived in Example 2.2.

Example 3.13. Let π be as in Example 2.3. Then

$$M = \begin{pmatrix} 2 & 5 & 4 & 1 & 1 & 0 \\ 3 & 2 & 4 & 0 & 0 & 2 \\ 5 & 3 & 2 & 2 & 5 & 0 \\ 0 & 1 & 1 & 4 & 3 & 2 \end{pmatrix},$$

$$K = \begin{pmatrix} 0 & -9 & 10 & 6 & -1 & 11 \\ 14 & 8 & -16 & 14 & -18 & 3 \end{pmatrix},$$

$$K' = \begin{pmatrix} 0 & -9 & 10 & 6 & -1 & 11 \\ 14 & -10 & 4 & 26 & -20 & -19 \end{pmatrix}.$$

Hence,

$$\begin{aligned} G_{\{2,5,6\}} &= \varphi(K'_{\{2,5,6\}}) \\ &= \{x_2^9 x_3^{10} x_4^6 x_5^{10} x_6^{11} - 1,\ x_1^{14} x_2^{10} x_3^4 x_4^{26} x_5^{20} x_6^{19} - 1\}. \end{aligned}$$

Calculating a Gröbner basis of $G_{\{2,5,6\}}$ with respect to lex order $x_6 > x_1 > \ldots > x_5$ and then applying T_6 we get the following set $G_{\{2,5\}}$:

$$\begin{aligned} \{&x_2^{61} x_3^{146} - x_1^{154} x_4^{172} x_5^{201}, -x_2^{23} x_3^{54} + x_1^{56} x_4^{62} x_5^{73}, \\ &-x_1^{98} x_4^{110} x_5^{128} + x_2^{38} x_3^{92} x_6, -x_1^{42} x_4^{28} x_5^{55} + x_2^{15} x_3^{38} x_6^2, \\ &-x_2^8 x_3^{16} + x_1^{14} x_4^{14} x_5^{18} x_6^3, -x_1^{28} x_4^{34} x_5^{37} + x_2^7 x_3^{22} x_6^5, \\ &-x_1^{14} x_2 x_4^{20} x_5^{19} + x_3^6 x_6^8, -x_2^9 x_3^{10} x_4^6 x_5 + x_6^{11}\}. \end{aligned}$$

In the same way we can calculate $G_{\{2\}}$, which has 22 elements, and finally $G_\varnothing$, with 14 elements. Calculating a Gröbner basis of $G_\varnothing$ with respect to lex order $x_1 > \cdots > x_6$ one can check that this result coincides with the one derived in Example 2.3.

4. SIMULATION RESULTS AND DISCUSSION

The proposed algorithm requires the determination of at most $\lfloor \frac{1}{2} n \rfloor$ Gröbner bases over K_x. Based on the (empirical) fact that the complexity of the Buchberger algorithm is a strongly growing function of the number of variables, we conclude that it is in general more efficient to evaluate $\lfloor \frac{1}{2}n \rfloor$ Gröbner bases over K_x than one Gröbner basis over $K_{x,y}$. In the special cases where either $\ker \pi_*$ (hence also $\ker \pi$) is trivial or $\ker \pi_*$ is spanned by a single element in $\mathbb{Z}^n$, no determination of a Gröbner basis is required in the proposed algorithm; more generally, we observe that the starting point of the proposed algorithm is a collection of binomials, whose cardinality is equal to the nullity of a generic linear map from $\mathbb{Z}^n$ to $\mathbb{Z}^m$: namely, $\max\{n-m, 0\}$.

Table 1 compares the running times of the standard versus the proposed algorithm for some parameters of n and m. For each pair (n, m) the time listed is the median of the running times, in seconds, of 500 random examples, where each entry in the matrix M was chosen independently according to a uniform distribution on the set $\{0, \dots, 5\}$. The simulation was performed on a NeXT computer (25 Mhz, 20 MBytes of RAM) using Mathematica [Wolfram 1991] (which uses lex order) to perform the Gröbner basis calculations and PARI [Batut et al. 1993] to perform the calculation of the integer kernel.

(n, m)	standard alg.	present alg.
$(3, 2)$	0.233	0.015
$(3, 3)$	0.383	0.015
$(4, 2)$	0.433	0.298
$(4, 3)$	1.350	0.015
$(4, 4)$	2.450	0.015
$(5, 2)$	0.883	0.864
$(5, 3)$	4.300	0.681
$(5, 4)$	11.333	0.031
$(5, 5)$	19.117	0.015
$(6, 2)$	1.450	2.046
$(6, 3)$	10.433	3.279
$(6, 4)$	43.283	1.448
$(6, 5)$	93.950	0.046
$(6, 6)$	190.233	0.031

TABLE 1. Comparison of running times between the standard and the present algorithms. Times listed are in seconds, and represent the median running times of 500 random examples where each entry of M is independent and uniformly distributed in the set $\{0, \dots, 5\}$.

From this table we see that for all listed cases but the case $(n, m) = (6, 2)$ the proposed algorithm performs better than the standard algorithm. As expected, the difference in performance becomes the more significant the larger m is compared to n. Beside the significant decrease in running time, the proposed algorithm also requires a significantly smaller amount of memory. For the mapping given in Example 2.3 the largest intermediate Gröbner basis required for the proposed algorithm has 22 elements, compared to 1180 elements that the standard algorithm requires.

The listed running times could be further reduced in several ways. Instead of using Mathematica to perform the Gröbner basis calculation, one could use Macaulay [Bayer and Stillman], which is significantly faster. More substantially, all occurring ideals are *toric ideals*, for which specialized Buchberger algorithms have been investigated [Conti and Traverso 1991; Hosten and Sturmfels 1994]. These speedups apply to the standard as well as to the proposed algorithm. There are at least two more potential ways in which the proposed algorithm can be made more efficient. First, it is known that the lexicographical ordering is in general not very efficient, so a potential improvement would be to replace lex order by a more efficient order which eliminates x_j. Secondly, in the above simulations no special effort was made to choose a specific K' from the many possible equivalent bases in order to minimize the subsequent calculations.

Although the original task of the proposed algorithm was to compute the kernel of a polynomial map, assuming that the map sends monomials into monomials, the algorithm also applies to ring homomorphisms $\pi : K_x \to K_{y,y^{-1}}$, where $K_{y,y^{-1}} := K[y_1, \dots, y_m, y_1^{-1}, \dots, y_m^{-1}]$. For these maps, the entries of the corresponding matrix M will be integers (rather than non negative integers), but the algorithm works as well.

Finally, we observe that in the standard method one could use the FGLM algorithm [Marinari et al. 1993], which given a Gröbner basis with respect to a certain ordering produces a basis with respect to another ordering. Since $\langle\{\gamma_1, \dots, \gamma_n\}\rangle$ is already a Gröbner basis (with respect to any term order that eliminates the x variables), one could apply this FGLM algorithm to $\langle\{\gamma_1, \dots, \gamma_n\}\rangle$ to compute a Gröbner basis with respect to a term order that eliminates the y variables, rather than computing this basis from scratch. But it seems that no substantial improvement in running times results from this approach. Further, this approach is limited to the original case where the map sends monomials into monomials, and does not extend to the more general case.

REFERENCES

[Adams and Loustaunau 1994] W. W. Adams and P. Loustaunau, *An Introduction to Gröbner Bases*, Graduate Studies in Math. **3**, Amer. Math. Soc., Providence, 1994.

[Batut et al. 1993] C. Batut, D. Bernardi, H. Cohen and M. Olivier, *User's Guide to Pari-GP*. This manual is part of the program distribution, available by anonymous ftp from the host megrez.ceremab.u-bordeaux.fr.

[Bayer and Stillman] D. Bayer and M. Stillman, *Macaulay: A Computer Algebra System for Algebraic Geometry*. This manual is part of the program distribution, available by anonymous ftp from the host available by anonymous ftp from zariski.harvard.edu.

[Buchberger 1985] B. Buchberger, "Gröbner bases: An algorithmic method in polynomial ideal theory", Chapter 6 in *Multidimensional Systems Theory* (edited by N. K. Bose), D. Reidel, Dordrecht and Boston, 1985.

[Cohen 1993] H. Cohen, *A Course in Computational Algebraic Number Theory*, Graduate Texts in Math. **138**, Springer, New York, 1993.

[Conti and Traverso 1991] P. Conti and C. Traverso, "Gröbner bases and integer programming", pp. 130–139 in *Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes*, New Orleans, 1991 (edited by H.F. Mattson et al.), Lecture Notes in Computer Science **539**, Springer, Berlin, 1991.

[Cox et al. 1991] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, Undergraduate Texts in Math., Springer, New York, 1991.

[Diaconis and Sturmfels] P. Diaconis and B. Sturmfels, "Algebraic algorithms for sampling from conditional distributions", to appear in *Ann. Stat.*

[Herzog 1970] J. Herzog, "Generators and relations of abelian semigroups and semigroup rings," *Manuscripta Math.* **3** (1970), 175–193.

[Hosten and Sturmfels 1994] S. Hosten and B. Sturmfels, "GRIN: An implementation of Gröbner bases for integer programming", 1994 (for information contact Hosten at serkan@math.berkeley.edu).

[Marinari et al. 1993] M. Marinari, H. Möller, and T. Mora, "Gröbner bases of ideals defined by functionals with an application to ideals of projective points", *Appl. Algebra Eng. Commun. Comput.* **4** (1993), 103–145.

[Natraj et al.] N. R. Natraj, S. R. Tayur and R. R. Thomas, "An algebraic geometry algorithm for scheduling in the presence of setups and correlated demands", to appear in *Math. Programming.*

[Thomas] R. R. Thomas, "A geometric Buchberger algorithm for integer programming", to appear in *Math. Oper. Res.*

[Wolfram 1991] S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, 2nd ed., Addison-Wesley, Reading, MA, 1991.

Fausto Di Biase, Department of Mathematics, Princeton University, Princeton, NJ 08544 (biase@math.princeton.edu)

Rüdiger Urbanke, Room 2C-254, AT&T Bell Labs, 600 Mountain Avenue, P.O.Box 636, Murray Hill, NJ 07974 (ruediger@research.att.com)

Received May 19, 1994; accepted in revised form July 21, 1995