J. Symbolic Computation (2001) **31**, 565–584
doi:10.1006/jsco.2001.0442
Available online at http://www.idealibrary.com on IDEAL®

Gröbner Bases over Galois Rings with an Application to Decoding Alternant Codes

EIMEAR BYRNE† AND PATRICK FITZPATRICK†

Department of Mathematics, National University of Ireland, Cork, Ireland

We develop a theory of Gröbner bases over Galois rings, following the usual formulation for Gröbner bases over finite fields. Our treatment includes a division algorithm, a characterization of Gröbner bases, and an extension of Buchberger's algorithm. One application is towards the problem of decoding alternant codes over Galois rings. To this end we consider the module $M = \{(a, b) : aS \equiv b \bmod x^r\}$ of all solutions to the so-called key equation for alternant codes, where S is a syndrome polynomial. In decoding, a particular solution $(\Sigma, \Omega) \in M$ is sought satisfying certain conditions, and such a solution can be found in a Gröbner basis of M. Applying techniques introduced in the first part of this paper, we give an algorithm which returns the required solution.

1. Introduction

The theory of Gröbner bases has been applied in several ways to error correcting codes. For example, in Fitzpatrick (1995) new algorithms corresponding to the Euclidean, Berlekamp–Massey, and Peterson–Gorenstein–Zierler algorithms for the solution of the key equation that arises in decoding alternant codes were derived from the perspective of Gröbner bases. Each of these algorithms is computationally at least as efficient as its classical analogue (Fitzpatrick, 1995; Fitzpatrick and Jennings, 1998; Fitzpatrick, 1999). The Gröbner basis approach has been extended to rational approximation and interpolation problems and to the solution of multivariable congruences (Fitzpatrick, 1996, 1997). Motivated by this research, we apply similar principles to the problem of decoding alternant codes defined over a Galois ring. The reader is referred to Shankar (1979), Hammons *et al.* (1994), Calderbank and Sloane (1995), Kanwar and Lopez-Permouth (1997) for review of the literature on the theory of codes over rings.

Whereas the classic theory of Gröbner bases assumes the coefficients lie in a field (see Cox *et al.*, 1992; Becker and Weispfenning, 1993; Adams and Loustaunau, 1994, for a review of the literature), in these investigations we extend the theory to the specific context of a Galois ring R. Many of our results are exact analogues of those holding over a field. However, their proofs are complicated by the change in significance of the coefficients, which may be zero divisors in R.

First we consider the notion of division in $R[\mathbf{x}]$, and generalize the division algorithm over a field to one for $R[\mathbf{x}]$. A central part of this task involves the idea of reducing the coefficients of a polynomial with respect to the coefficients of the division set and exploits the canonical representations of elements in R. We next establish the existence

†E-mail: {eimear.byrne,p.fitzpatrick}@ucc.ie

0747–7171/01/050565 + 20 $35.00/0

of Gröbner bases for arbitrary non-zero ideals in $R[\mathbf{x}]$, and characterize such bases. Since it is essential for our purposes to be able to compute a Gröbner basis from any given generating set, we develop an analogue of Buchberger's algorithm, and give conditions under which a generating set is indeed a Gröbner basis.

In the final part of the paper we summarize results presented in Byrne and Fitzpatrick (preprint) which give an application of the principles derived in Section 2 to the decoding problem. The procedure used is locator decoding, and is dependent on finding a particular solution of a key equation. It turns out that the required solution is minimal in a subset of the module of all solutions to the key equation (under a certain monomial ordering) and is thus contained in a Gröbner basis for the module. Using a known basis, and the extension of Buchberger's algorithm presented in Section 2.5 we compute the desired Gröbner basis.

We conclude this Introduction by recalling some of the basic properties of Galois rings. These have been well documented in McDonald (1974) and Raghavendran (1969), the former giving an explicit development along the lines of the theory of Galois fields. There are a number of equivalent descriptions of Galois rings: they are the *separable* extensions of finite, unital, local, commutative rings and the *unramified* extensions of such rings. Henceforth we assume that all rings R and T are finite, local, commutative rings with unity. We refer to an element of an arbitrary ring as regular if it is not a zero divisor in that ring.

Let T have maximal ideal $\langle p\rangle$ for some prime p. The polynomial $f \in T[x]$ is called a *basic irreducible* if it is irreducible modulo p. We construct the Galois ring as a quotient ring of $\mathbf{Z}_{p^n}[x]$ as follows. Let m, n be positive integers and let $f \in \mathbf{Z}_{p^n}[x]$ be a monic basic irreducible polynomial of degree m. The quotient ring $\mathbf{Z}_{p^n}[x]/\langle f\rangle$, denoted $\mathrm{GR}(p^{mn}, p^n)$, is called the *Galois ring* of order p^{mn} and characteristic p^n.

The integers p, m and n chosen as above determine uniquely (up to isomorphism) the Galois ring $\mathrm{GR}(p^{mn}, p^n)$ (Raghavendran, 1969, p. 207). For the remainder of the text the symbol R will denote a Galois ring, R^* its multiplicative group of units, k_R the residue field of R, and μ the natural epimorphism defined by

$$\mu : R \to k_R : a \mapsto a + \langle p\rangle.$$

The residue field of the ring R is unique and isomorphic to the finite field $\mathrm{GF}(p^m)$ of p^m elements.

Let $\bar{f}$ be a monic irreducible divisor of $x^N - 1$ in $\mathbf{Z}_p[x]$ of degree m where $N = p^m - 1$. Then Hensel's Lemma (McDonald, 1974, p. 256, Theorem XIII.4), ensures the existence of a unique monic irreducible $f \in \mathbf{Z}_{p^n}[x]$ such that f divides $x^N - 1$ in $\mathbf{Z}_{p^n}[x]$ and $\mu f = \bar{f}$. There is thus a one-to-one correspondence between the irreducible factors of $x^N - 1$ modulo p and the irreducible factors of $x^N - 1$ modulo p^n. If f is a primitive basic irreducible, and ξ is a root of f, then $k_R = \mathbf{Z}_p[\mu\xi]$ and $R = \mathbf{Z}_{p^n}[\xi]$.

Let $\mathcal{T}$ be a transversal on the cosets of $\langle p\rangle$ in R, so that if $\upsilon, \rho \in \mathcal{T}$ then $\upsilon - \rho \in \langle p\rangle$ if and only if $\upsilon = \rho$. There are two ways of uniquely representing an element θ in R. The first expression comes from adjoining the root ξ to the ring $\mathbf{Z}_{p^n}$, as illustrated below.

$$\theta = \sum_{j=0}^{m-1} a_j \xi^j, \qquad a_j \in \mathbf{Z}_{p^n}. \tag{1}$$

We refer to this representation of an element θ of R as the *additive normal form* of θ, since it is preserved under component-wise addition. Given a specified transversal $\mathcal{T}$, the

second type has the form

$$\theta = \sum_{j=0}^{n-1} p^j \theta_j, \qquad \theta_j \in \mathcal{T}, \tag{2}$$

which we call the *p-adic representation* of θ with respect to the transversal $\mathcal{T}$ (or simply the p-adic representation where it is assumed that an arbitrary transversal has been selected). That an element θ has this latter unique representation can be proved by an inductive argument. For a given $\theta \in R$, the element θ_j (or $(\theta)_j$ if parentheses are required to avoid ambiguity) is the uniquely determined jth component of θ in $\mathcal{T}$. Note that the p-adic representation is not preserved under addition.

2. Gröbner Bases in $R[\mathbf{x}]$

2.1. NOTATION

A *term* in $R[\mathbf{x}]$ is an element of the form $x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m}$ for some integers $i_j \geq 0$. In general, an arbitrary term will be denoted by X. A *monomial* in $R[\mathbf{x}]$ is a non-zero constant multiple of a term in $R[\mathbf{x}]$. Throughout this paper $<$ denotes an arbitrary fixed term order. Let $a \in R[\mathbf{x}]$. The expressions *leading term*, $\mathrm{lt}(a)$, *leading coefficient*, $\mathrm{lc}(a)$, and *leading monomial*, $\mathrm{lm}(a) = \mathrm{lc}(a)\mathrm{lt}(a)$ have the usual meanings.

Given any non-empty subset $\mathcal{S}$ of $R[\mathbf{x}]$, we denote by $\mathrm{lt}(\mathcal{S})$ (respectively $\mathrm{lm}(\mathcal{S})$) the set of leading terms (respectively monomials) of the elements of $\mathcal{S}$. The ideal of $R[\mathbf{x}]$ generated by the elements of $\mathcal{S}$ is denoted $\langle \mathcal{S} \rangle$ and we write $\mathrm{Lt}(\mathcal{S})$ for $\langle \mathrm{lt}(\mathcal{S}) \rangle$ and $\mathrm{Lm}(\mathcal{S})$ for $\langle \mathrm{lm}(\mathcal{S}) \rangle$.

2.2. DIVISION IN R

Central to the task of describing our division in $R[\mathbf{x}]$ is a description of a division process in R. In the presence of the zero divisors in R, we generalize the notion of division of one element by another to that of *reduction*.

Let $\theta \in R \backslash \{0\}$. Define the map

$$\nu_p : R\backslash\{0\} \longrightarrow \mathbf{N}_0 : \theta \longmapsto \min\{t : \theta \in \langle p^t \rangle\}.$$

Then $\nu_p(\theta)$ denotes the highest power of p which divides θ ($\theta = \beta p^{\nu_p(\theta)}$ for some $\beta \in R^*$), and $\mathrm{ann}(\theta) = \langle p^{n-\nu_p(\theta)} \rangle$. We extend this to a map on R by setting $\nu_p(0) = -1$.

Let $\theta \in R$ and let $\mathcal{T}$ be a transversal on the cosets of $\langle p \rangle$ in R. It has already been noted that θ can be expressed uniquely in its additive normal form and p-adically by Equations (1) and (2). For $1 \leq k \leq n$, denote by $\theta^{(k)}$ the truncation of θ modulo p^{k+1}, with respect to its p-adic representation:

$$\theta^{(k)} = \sum_{i=0}^{k} \theta_i p^i.$$

For $1 \leq k \leq n$, denote by $\theta^{|k|}$ the sum

$$\theta^{|k|} = \sum_{j=0}^{m-1} \bar{a}_j \xi^j$$

where $\bar{a}_j$ is the unique element of $\{0,\dots,p^k-1\}$ satisfying the relation $a_j \equiv \bar{a}_j \bmod p^k$. We also define the following maps:

$$\eta_p : R\backslash\{0\} \to \mathbf{N}_0 : \theta = \sum_{i=0}^{n-1} \theta_i p^i \longmapsto \max\{i : \theta_i \neq 0\}$$

$$\kappa_p : R\backslash\{0\} \to \mathbf{N}_0 : \theta = \sum_{j=0}^{m-1} a_j \xi^j \longmapsto \max\{\eta_p(a_j) : a_j \neq 0\}.$$

Note that if $\nu_p(\theta) = k$ for some non-negative integer k, then $\theta = \theta_k p^k$ if and only if $\eta_p(\theta) = k$, and $\theta = \theta_k p^k$ if and only if $\kappa_p(\theta) = k$. We extend these to mappings defined on R by setting $\eta_p(0) = \kappa_p(0) = n$.

We now introduce the notion of reduction in R.

Definition 2.2.1. Let $\theta, \lambda, \rho \in R$, with $\lambda \neq 0$. We say that θ *reduces to* ρ *modulo* λ *with respect to the p-adic representation* if there exists $\upsilon \in R$ such that

$$\theta = \upsilon\lambda + \rho$$

where $\eta_p(\rho) < \nu_p(\lambda)$ or $\rho = 0$. We say that θ is reduced modulo λ with respect to the p-adic representation if the only solution of this equation satisfying $\eta_p(\rho) < \nu_p(\lambda)$ is given by $\upsilon = 0$, $\theta = \rho$. Thus θ is reduced modulo λ with respect to the p-adic normal form if and only if $\eta_p(\theta) < \nu_p(\lambda)$.

We say that θ *reduces to* ρ' *modulo* λ *with respect to the additive normal form* if there exists $\upsilon' \in R$ such that

$$\theta = \upsilon'\lambda + \rho'$$

where $\kappa_p(\rho') < \nu_p(\lambda)$ or $\rho' = 0$. We say that θ is reduced modulo λ with respect to the additive normal form if the only solution of this equation satisfying $\kappa_p(\rho') < \nu_p(\lambda)$ is given by $\upsilon' = 0$, $\theta = \rho'$. Thus θ is reduced modulo λ with respect to the additive normal form if and only if $\kappa_p(\theta) < \nu_p(\lambda)$.

Lemma 2.2.2. *Let $\theta, \lambda \in R, \lambda \neq 0$. Then there exist unique ρ, ρ' such that:*

(i) θ reduces to ρ modulo λ with respect to a p-adic representation;
(ii) θ reduces to ρ' modulo λ with respect to the additive normal form.

Proof. The proof is outlined in the reduction process. For $\theta, \lambda \in R$, with $\lambda \neq 0$, write

$$\theta = \sum_{i=0}^{n-1} \theta_i p^i, \qquad \lambda = \beta p^k$$

where $\theta_i \in \mathcal{T}$ for some transversal $\mathcal{T}$ on the cosets of $\langle p \rangle$ in R, $\nu_p(\lambda) = k$ and $\beta \in R^*$. Then we write

$$\begin{aligned}\theta &= \left(\sum_{i=k}^{n-1} \theta_i p^{i-k}\right)\beta^{-1}\lambda + \theta^{(k-1)} \\ &= \upsilon\lambda + \rho\end{aligned}$$

where $\upsilon = \beta^{-1}\big(\sum_{i=k}^{n-1} \theta_i p^{i-k}\big)$ and $\rho = \theta^{(k-1)} = \sum_{i=0}^{k-1} \theta_i p^i$, as defined earlier. Then $\eta_p(\rho) < \nu_p(\lambda)$ unless $\rho = 0$, and it is clear that θ reduces to ρ uniquely. This proves (i).

Similarly, expressing θ in additive normal form we have

$$\theta = \sum_{j=0}^{m-1} a_j \xi^j, \qquad \lambda = \beta p^k$$

where $\nu_p(\lambda) = k$. Then

$$\begin{aligned}\theta &= \left(\sum_{j=0}^{m-1} a'_j \xi^j\right)\beta^{-1}\lambda + \theta^{|k|} \\ &= \upsilon'\lambda + \rho'\end{aligned}$$

where $\upsilon' = \left(\sum_{j=0}^{m-1} a'_j \xi^j\right)\beta^{-1}$, and $\rho' = \theta^{|k|} = \sum_{j=0}^{m-1} \bar{a}_j \xi^j$, where for each $j \in \{0, \ldots, m-1\}$, $a_j = a'_j p^k + \bar{a}_j$ for some $a'_j \in \mathbf{Z}_{p^n}$ and $\bar{a}_j \in \{0, \ldots, p^k - 1\}$ satisfying $a_j \equiv \bar{a}_j \bmod p^k$. Then $\kappa_p(\rho') < \nu_p(\lambda)$ unless $\rho' = 0$, and ρ' is uniquely determined by this reduction process, proving (ii). □

Remark 2.2.3. The remainder produced in the reduction of θ modulo λ, is simply $\theta^{(k-1)}$ or $\theta^{|k|}$, depending on the procedure chosen, and depends only on $k = \nu_p(\lambda)$. When λ is a unit both reduction procedures are identical to ordinary division. Note, however, that in general these two reduction procedures do not lead to the same remainder, since for any $\alpha \in R$, the values $\eta_p(\alpha)$ and $\kappa_p(\alpha)$ do not necessarily coincide. For example, in $\mathrm{GR}(8^3, 8) \simeq Z_8[\xi]$ the element ξ^6 is in p-adic representation with respect to the transversal $\mathcal{T} = \{0, 1, \xi, \ldots, \xi^6\}$, and is expressed in additive normal form as $5 + 6\xi + \xi^2$. However, $\eta_p(\xi^6) = 0$, while $\kappa_p(\xi^6) = 2$. According to the first reduction process, ξ^6 is reduced modulo 2, while in the latter, we find that $5 + 6\xi + \xi^2$ reduces to $1 + \xi^2$ modulo 2. On some occasions one reduction procedure may be more convenient to use than the other; what is vital is that the process selected must be adhered to for the duration of the calculations. Henceforth, any statements made regarding division where we do not specify which reduction procedure is being implemented can be assumed to be true for the exclusive use of either procedure.

An exception to this general rule is illustrated by the following. Let $R = \mathbf{Z}_9$. Consider the element 7. It has additive normal form 7, and is expressed in p-adic representation as $1 + (2)(3)$, with respect to the transversal $\{0, 1, 2\}$. Now 7 reduces to 1 modulo 3 with respect to the additive normal form, and 7 reduces to 1 modulo 3 with respect to this p-adic representation, so in this case both reduction procedures lead to the same remainder. In general, if $R = \mathbf{Z}_{p^n}$ then reduction with respect to the additive normal form is identical to reduction with respect to the p-adic representation when the transversal chosen is $\{0, 1, \ldots, p-1\}$. We state this as follows.

Lemma 2.2.4. *Let $R = \mathbf{Z}_{p^n}$. Let $\theta, \lambda \in R$ such that $\lambda \neq 0$. Then there exists $\rho \in R$ such that θ reduces to ρ modulo λ with respect to the additive normal form, and θ reduces to ρ modulo λ with respect to the p-adic representation when the transversal chosen is $\{0, 1, \ldots, p-1\}$.*

Proof. We show for $\theta \in R$, that $\theta^{(k)} = \theta^{|k+1|}$ for all $k \in \{0, \ldots, n\}$ when $R = \mathbf{Z}_{p^n}$. In p-adic notation θ is given by $\theta_0 + p\theta_1 + \cdots + p^{n-1}\theta_{n-1}$, with each $\theta_i \in \{0, \ldots, p-1\}$.

Now $\theta^{|k+1|} = \bar{\theta}$ where $\bar{\theta}$ is the unique element of $\{0, \ldots, p^{k+1} - 1\}$ such that $\theta \equiv \bar{\theta} \bmod p^{k+1}$, and $\theta^{(k)} = \theta_0 + \cdots + p^k\theta_k \in \{0, \ldots, p^{k+1} - 1\}$. It follows that $\theta^{(k)} = \theta^{|k+1|}$. □

LEMMA 2.2.5. *Let $\theta = \lambda_1 + \cdots + \lambda_s$ where $\theta, \lambda_j \in R$ and $\lambda_j \neq 0$ for any j. Then θ reduces to 0 modulo λ_j (i.e. θ is divisible by λ_j) for at least one j.*

PROOF. Let $j \in \{1, \ldots, s\}$ such that $\nu_p(\lambda_j) = \min\{\nu_p(\lambda_i) : i \in \{1, \ldots, s\}\}$, and let $\nu_p(\theta) = t$. Then $k \leq t$, and writing $\lambda_j = \beta p^k, \theta = \gamma p^t$ for some $\beta, \gamma \in R^*$, we have that $\theta = (\gamma\beta^{-1}p^{t-k})\lambda_j$. □

In other words an element of R is a sum of a set of non-zero elements in R if and only if it is divisible by at least one element in the set. In particular it must be divisible by the element λ with $\nu_p(\lambda)$ minimal.

LEMMA 2.2.6. *Let $\theta, \theta', \lambda \in R$ with $\lambda \neq 0$. Then θ and θ' have the same remainder with respect to the p-adic representation if and only if θ and θ' have the same remainder with respect to the additive normal form.*

PROOF. Let $\nu_p(\lambda) = k$ for some integer k. Then $\theta^{(k-1)} = \theta'^{(k-1)}$ if and only if $\theta + \langle p^k \rangle = \theta' + \langle p^k \rangle$ if and only if $\theta^{|k|} = \theta'^{|k|}$. The result follows. □

LEMMA 2.2.7. *Let $\nu_p(\lambda) = k$. Then $\theta + \langle p^t \rangle = \theta' + \langle p^t \rangle$ for some $t \geq k$, if and only if θ and θ' both have the same remainder modulo λ.*

PROOF. Suppose $\theta + \langle p^t \rangle = \theta' + \langle p^t \rangle$ for some $t \geq k$. Then $\theta^{(t-1)} = \theta'^{(t-1)}$, and hence $\theta^{(l)} = \theta'^{(l)}$ for all $l \in \{0, \ldots, t-1\}$. In particular $\theta^{(k-1)} = \theta'^{(k-1)}$ and $\theta^{|k|} = \theta'^{|k|}$ since $k \leq t$, so that θ and θ' both have the same remainder modulo λ. Conversely, suppose that θ and θ' both have the same remainder modulo λ. Then $\theta^{(k-1)} = \theta'^{(k-1)}$, and $\theta^{|k|} = \theta'^{|k|}$. Set $t = k$. Then $\theta + \langle p^t \rangle = \theta' + \langle p^t \rangle$ and the result is proved. □

2.3. DIVISION IN $R[\mathbf{x}]$

We introduce a division algorithm in $R[\mathbf{x}]$. We emphasize that since in general p-adic representation reduction will lead to a different remainder from the additive normal form reduction, the algorithm admits the use of only one pre-selected reduction process.

DEFINITION 2.3.1. Let a, b, r be monomials in $R[\mathbf{x}]$ with $b \neq 0$. We say that a *reduces to r modulo b* if there exists $\theta \in R$, and a term $X \in R[\mathbf{x}]$, such that

$$a = \theta X b + r$$
$$\mathrm{lt}(a) = \max\{X\mathrm{lt}(b),\ \mathrm{lt}(r)\}$$

and $\mathrm{lt}(r) = X\mathrm{lt}(b)$ only if $\mathrm{lc}(r)$ is reduced modulo $\mathrm{lc}(b)$. The monomial a is said to be *reduced* modulo b if the only solution satisfying the given criteria is $\theta = 0$ and $r = a$. Otherwise we say that a is *reducible* modulo b.

DEFINITION 2.3.2. Let $a, b, r \in R[\mathbf{x}]$ with $b \neq 0$. We say that a *reduces to r in one step modulo b*, denoted $a \to_b r$, if there exists a monomial a_1 of a that reduces to a monomial

r_1 of r modulo $\mathrm{lm}(b)$, that is, if there exist $\theta \in R$, and $X \in R[\mathbf{x}]$, such that

$$a = \theta X b + r$$
$$a_1 = \theta X \mathrm{lm}(b) + r_1$$
$$\mathrm{lt}(a) = \max\{X\mathrm{lt}(b), \mathrm{lt}(r)\}$$

where $X\mathrm{lt}(b)$ occurs as a term of r only if the coefficient of that term is reduced modulo $\mathrm{lc}(b)$. The element *a is reduced modulo b* if every monomial of a is reduced modulo $\mathrm{lm}(b)$.

The division algorithm now follows in the usual way. Let $a, r \in R[\mathbf{x}]$, $\mathcal{G} = \{b_i\}_1^s \subseteq R[\mathbf{x}]$ with $\mathcal{G} \neq \{0\}$. We write $a \to_{+\mathcal{G}} r$ if r can be obtained by a sequence of one-step reductions. If $a \to_{+\mathcal{G}} r$, and r is reduced modulo $\mathcal{G}$, then r is called a remainder of a with respect to $\mathcal{G}$, and we write $\mathrm{rem}_{\mathcal{G}}(a)$. Note that $\mathrm{rem}_{\mathcal{G}}(a)$ is not unique for an arbitrary set $\mathcal{G}$. The division algorithm is stated as follows.

THEOREM 2.3.3. *Given $a \in R[\mathbf{x}], \mathcal{G} = \{b_i\}_1^s \subseteq R[\mathbf{x}]$, there exist $\{f_i\}_1^s \subseteq R[\mathbf{x}]$, $r \in R[\mathbf{x}]$ satisfying*

$$a = \sum_1^s f_i b_i + r$$

where $\mathrm{lt}(a) = \max\{\mathrm{lt}(f_i)\mathrm{lt}(b_i), \mathrm{lt}(r)\}$ and either each monomial of r is reduced modulo $\mathcal{G}$ or $r = 0$.

2.4. CHARACTERIZATION OF GRÖBNER BASES IN $R[\boldsymbol{x}]$

DEFINITION 2.4.1. Let A be an ideal in $R[\mathbf{x}]$. A set $\mathcal{G} = \{b_i\}_1^s \subseteq A$ of non-zero elements is called a *Gröbner basis* of A if for each $a \in A$ there exists an $i \in \{1, \ldots, s\}$ such that $\mathrm{lm}(a)$ is divisible by $\mathrm{lm}(b_i)$. An arbitrary subset $\mathcal{G}$ of $R[\mathbf{x}]$ is called a Gröbner basis if it is a Gröbner basis of $\langle \mathcal{G} \rangle$.

From Lemma 2.2.5 it follows that if a monomial a can be reduced to zero by the action of some non-empty set of monomials $\mathcal{G}$ then a must be divisible by some member of $\mathcal{G}$, in other words:

LEMMA 2.4.2. *Let a, $\mathcal{G} = \{b_i\}_1^s$ be monomials in $R[\mathbf{x}]$. Suppose that $a \to_{+\mathcal{G}} 0$. Then there exists $i \in \{1, \ldots, s\}$ such that $a \to_{b_i} 0$.*

The following theorem gives the expected characterization of a Gröbner basis of an ideal in $R[\mathbf{x}]$.

THEOREM 2.4.3. *Let A be an ideal in $R[\mathbf{x}]$, $\mathcal{G} = \{b_i\}_1^s \subseteq A$. The following statements are equivalent.*

(i) $\mathcal{G}$ *is a* Gröbner basis *of* A.
(ii) $a \in A$ *if and only if* $a \to_{+\mathcal{G}} 0$.
(iii) For all $a \in A$ *there exist* $\{f_i\}_1^s \subset R[\mathbf{x}]$ *satisfying*

$$a = \sum f_i b_i$$
$$\mathrm{lt}(a) = \max\{\mathrm{lt}(f_i)\mathrm{lt}(b_i)\}.$$

(iv) $\mathrm{Lm}(\mathcal{G}) = \mathrm{Lm}(A)$.

Clearly if $\mathcal{G}$ is a Gröbner basis of an ideal A then $\langle \mathcal{G} \rangle = A$.

Since our definition of reduction of a polynomial in $R[\mathbf{x}]$ permits us to reduce a polynomial a by some set $\mathcal{G}$ by reducing any monomial of a, we can implement a straightforward adaptation of (Adams and Loustaunau, 1994, Corollary 1.8.6), to generate a *reduced* Gröbner basis from a minimal Gröbner basis of an ideal in $R[\mathbf{x}]$. It is easy to show that any ideal in $R[\mathbf{x}]$ has a unique reduced Gröbner basis.

We have already mentioned that for an arbitrary set $\mathcal{G}$ the division algorithm may not yield uniquely defined remainders. However, in the case where $\mathcal{G}$ is a Gröbner basis then the remainder found on division of a by $\mathcal{G}$ is indeed unique.

THEOREM 2.4.4. *Let $\mathcal{G} = \{b_i\}_1^s$. If $\mathcal{G}$ is a Gröbner basis then $a \rightarrow_{+\mathcal{G}} r$ uniquely for all $a \in R[\mathbf{x}]$.*

PROOF. Suppose that two different sequences of reductions yield:

$$\begin{aligned} a &= \sum_1^s f_i b_i + r \\ &= \sum_1^s f_i' b_i + r' \end{aligned}$$

where

$$\begin{aligned} \mathrm{lt}(a) &= \max\{\mathrm{lt}(f_i)\mathrm{lt}(b_i), \mathrm{lt}(r)\} \\ &= \max\{\mathrm{lt}(f_i')\mathrm{lt}(b_i), \mathrm{lt}(r')\} \end{aligned}$$

and each monomial of r and r' is reduced modulo b_i for $i = 1, \ldots, s$. Then

$$r - r' = \sum_1^s (f_i' - f_i) b_i \in \langle \mathcal{G} \rangle.$$

Since $\mathcal{G}$ is a Gröbner basis, it follows that $\mathrm{lm}(r - r')$ is divisible by $\mathrm{lm}(b_i)$ for some $i \in \{1, \ldots, s\}$ unless $r - r' = 0$. Now

$$\mathrm{lt}(r - r') \leq \max\{\mathrm{lt}(r), \mathrm{lt}(r')\}$$

and if $\mathrm{lt}(r) > \mathrm{lt}(r')$ then $\mathrm{lm}(r - r') = \mathrm{lm}(r)$ which is not divisible by the leading term of any element of $\mathcal{G}$ so that a contradiction results. It follows that $\mathrm{lt}(r) = \mathrm{lt}(r')$. If $\mathrm{lc}(r) \neq \mathrm{lc}(r')$ then

$$\begin{aligned} \mathrm{lm}(r - r') &= \mathrm{lm}(r) - \mathrm{lm}(r') \\ &= (\mathrm{lc}(r) - \mathrm{lc}(r'))\mathrm{lt}(r). \end{aligned}$$

By hypothesis, the monomials of both r and r' are reduced with respect to the leading monomial of each element of $\mathcal{G}$. In particular, if $\mathrm{lt}(b_i)$ divides $\mathrm{lt}(r)$ for some $i \in \{1, \ldots, s\}$ then $\mathrm{lc}(r)$ and $\mathrm{lc}(r')$ must be reduced modulo $\mathrm{lc}(b_i)$. Then $\mathrm{lc}(r) \neq \mathrm{lc}(r')$ implies that $\mathrm{lc}(r - r') = \mathrm{lc}(r) - \mathrm{lc}(r')$ is divisible by $\mathrm{lc}(b_j)$ for some $j \in \{1, \ldots, s\}$, and therefore $\mathrm{lc}(r) - \mathrm{lc}(r') \in \langle p^l \rangle$ where $\mathrm{lc}(b_j) = \beta p^l$ for some $\beta \in R^*$, $l \geq 1$. But then by Lemma 2.2.7, $\mathrm{lc}(r)$ and $\mathrm{lc}(r')$ both have the same remainder modulo b_j. Since $\mathrm{lc}(r)$ and $\mathrm{lc}(r')$ are both reduced modulo $\mathrm{lc}(b_j)$, it follows that $\mathrm{lc}(r) = \mathrm{lc}(r')$ and thus $\mathrm{lm}(r) = \mathrm{lm}(r')$. Repeated applications of the same argument yield that r and r' agree at each coefficient and the remainder is unique. □

DEFINITION 2.4.5. Let $\mathcal{G}$ be a Gröbner basis and let $f \in R[\mathbf{x}]$. The *normal form of* f *with respect to* $\mathcal{G}$, denoted $\mathrm{Nf}_{\mathcal{G}}(f)$, is the unique remainder found in the application of the division algorithm to f and $\mathcal{G}$.

Theorem 2.4.4 guarantees the existence of a well defined *normal form*, with respect to a given Gröbner basis $\mathcal{G}$. The set of normal forms with respect to $\mathcal{G}$ forms a complete set of coset representatives for $R[x]/\langle\mathcal{G}\rangle$.

EXAMPLE 2.4.6. Let $R[\mathbf{x}] = \mathbf{Z}_{27}[x, y]$ and let $\mathcal{G} = \{g_j\}_{j=1}^{4}$ where g_1, g_2, g_3, g_4 are the polynomials $9, x + 1, 3y^2, y^3 + 13y^2 - 12$, *respectively. Endowing the terms of* $R[\mathbf{x}]$ with graded lex order with $y < x$ *gives*

$$1 < y < x < y^2 < xy < x^2 < y^3 < xy^2 < \cdots.$$

Let $\mathcal{F} = \{f_1, f_2\}$ where $f_1 = x^5y^2 + 2y^3 + 3x^2 + 6x + 6$ and $f_2 = 3y^2 + x + 1$. Then

$$f_1 = (x^4y^2 - x^3y^2 + x^2y^2 - xy^2 + y^2 + 3x - 3)g_2 + 2g_4, \quad \text{and}$$
$$f_2 = g_3 + g_2,$$

so that $\mathcal{G}$ generates $\langle\mathcal{F}\rangle$. In fact it can be shown (see Example 2.5.11) that $\mathcal{G}$ is a Gröbner basis of $\langle\mathcal{F}\rangle$. Let $f = x^6+x^5-9x^2+2xy+9x+2y+9$. Then $f = (x^5+2y)f_2-3f_1 \in \langle\mathcal{F}\rangle$, but applying the division algorithm to f and $\mathcal{F}$, we find that $f \to_{+\mathcal{F}} f$, since each monomial of f is reduced modulo $\mathrm{lm}(f_1)$ and $\mathrm{lm}(f_2)$. However, $f \to_{+\mathcal{G}} 0$ via the reduction

$$f \to_{g_2} 18x^2 + 2xy + 9x + 2y + 9 \to_{g_2} 2xy + 18x + 2y + 9 \to_{g_2} 18x + 9 \to_{+g_1} 0$$

so that

$$f = (x^5 - 9x + 2y)g_2 + (2x + 1)g_1$$

and $\mathrm{lt}(f) = x^6 = \mathrm{lt}(x^5 - 9x + 2y)\ \mathrm{lt}(g_2)$, as predicted by Theorem 2.4.3. Since $\mathcal{G}$ is a Gröbner basis, the remainder produced in the application of the division algorithm to $\mathcal{G}$ and an element of $R[\mathbf{x}]$ must be unique. Consider the following reduction:

$$f \to_{g_2} 18x^2 + 2xy + 9x + 2y + 9 \to_{+g_1} 2xy + 2y \to_{g_2} 0.$$

So

$$f = (x^5 + 2y)g_2 + (x^2 + 1)g_2$$

and $\mathrm{lt}(f) = x^6 = \mathrm{lt}(x^5 + 2y)\mathrm{lt}(g_2)$. Although the sequence of reductions differs, we still get a zero remainder. We can always determine whether or not an element f is contained in $\langle\mathcal{F}\rangle$ by checking its remainder on division by the set $\mathcal{G}$.

2.5. COMPUTING GRÖBNER BASES IN $R[\mathbf{x}]$

Having established the existence of a Gröbner basis for every non-zero ideal in $R[\mathbf{x}]$ and devised a division algorithm, it remains to show how to compute a Gröbner basis for an ideal from a given generating set. Our approach is essentially a generalization of Buchberger's algorithm. We denote by $\mathbf{e}_i$ the vector with 1 in position i and 0 elsewhere (and length implied by the context). Our algorithm is based on the following theorem (Adams and Loustaunau, 1994, p.213, Theorem 4.2.3).

THEOREM 2.5.1. *Let $\mathcal{G} = \{g_i\}_{i=1}^s$ be a set of non-zero polynomials in $R[\mathbf{x}]$. Let $\mathcal{B}$ be a homogeneous generating set for* $\mathrm{Syz}(lm(g_1), \ldots, lm(g_s))$. *Then $\mathcal{G}$ is a Gröbner basis for $\langle \mathcal{G} \rangle$ if and only if for all $[h_1, \ldots, h_s] \in \mathcal{B}$*

$$h_1 g_1 + \cdots + h_s g_s \longrightarrow_{+\mathcal{G}} 0.$$

In what follows we find a homogeneous generating set for the syzygy module of an ordered s-tuple of monomials in $R[\mathbf{x}]$, and hence develop an appropriate algorithm for the computation of a Gröbner basis in $R[\mathbf{x}]$. In particular, we shall prove the following theorem.

THEOREM 2.5.2. *Let $\upsilon_i X_i, i = 1, \ldots, s$ be monomials in $R[\mathbf{x}]$. Express υ_i as $\upsilon_i = \theta_i p^{t_i}$ for some $\theta_i \in R^*, t_i \geq 0$. Then the syzygy module* $\mathrm{Syz}(\upsilon_1 X_1, \ldots, \upsilon_s X_s)$ *is generated by*

$$\{p^{n-t_i}\mathbf{e}_i : i = 1, \ldots, s\} \cup \left\{ p^{t_{i,j}-t_i}\theta_i^{-1}\frac{X_{ij}}{X_i}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\frac{X_{ij}}{X_j}\mathbf{e}_j : 1 \leq i < j \leq s \right\}$$

where $t_{i,j} = \max\{t_i, t_j\}, X_{ij} = \mathrm{lcm}\{X_i, X_j\}$.

Before proving this result we consider the nature of generating sets of the related syzygy module $\mathrm{Syz}(\upsilon_1, \ldots, \upsilon_s) \subseteq R^s$.

DEFINITION 2.5.3. Let $\upsilon_i = \theta_i p^{t_i}$ as above. Let $\lambda = (\lambda_1, \ldots, \lambda_s) \in \mathrm{Syz}(\upsilon_1, \ldots, \upsilon_s)$ with $\lambda_i = \beta_i p^{k_i}, \beta_i \in R^* \cup \{0\}$. We say that λ is homogeneous of degree p^l if $k_i + t_i = l$ for each $i = 1, \ldots, s$ for which $\beta_i \neq 0$.

LEMMA 2.5.4. *With the notation above, let $\lambda \in \mathrm{Syz}(\upsilon_1, \ldots, \upsilon_s)$ be homogeneous of degree p^l. Then λ can be expressed as a linear combination of*

$$\{p^{n-t_i}\mathbf{e}_i : i = 1, \ldots, s\} \cup \{p^{t_{i,j}-t_i}\theta_i^{-1}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\mathbf{e}_j : 1 \leq i < j \leq s\}$$

where $t_{i,j} = \max\{t_i, t_j\}$.

PROOF. By definition $p^l = p^{k_i+t_i}$ for $i = 1, \ldots, s$. Thus

$$\begin{aligned}
\lambda &= (\beta_1 p^{k_1}, \ldots, \beta_s p^{k_s}) \\
&= \sum_{i=1}^{s} \beta_i p^{k_i}\mathbf{e}_i = \sum_{i=1}^{s} \beta_i\theta_i p^{l-t_i}\theta_i^{-1}\mathbf{e}_i \\
&= \beta_1\theta_1 p^{l-t_{1,2}}(p^{t_{1,2}-t_1}\theta_1^{-1}\mathbf{e}_1 - p^{t_{1,2}-t_2}\theta_2^{-1}\mathbf{e}_2) \\
&\quad +(\beta_1\theta_1 + \beta_2\theta_2)p^{l-t_{2,3}}(p^{t_{2,3}-t_2}\theta_2^{-1}\mathbf{e}_2 - p^{t_{2,3}-t_3}\theta_3^{-1}\mathbf{e}_3) + \cdots \\
&\quad +(\beta_1\theta_1 + \cdots + \beta_{s-1}\theta_{s-1})p^{l-t_{s-1,s}}(p^{t_{s-1,s}-t_{s-1}}\theta_{s-1}^{-1}\mathbf{e}_{s-1} - p^{t_{s-1,s}-t_s}\theta_s^{-1}\mathbf{e}_s) \\
&\quad +(\beta_1\theta_1 + \cdots + \beta_s\theta_s)p^{l-t_s}\theta_s^{-1}\mathbf{e}_s.
\end{aligned}$$

Since $\beta_1\theta_1 + \cdots + \beta_s\theta_s \in \langle p^{n-l} \rangle$ it follows that $(\beta_1\theta_1 + \cdots + \beta_s\theta_s)p^{l-t_s}\theta_s^{-1} \in \langle p^{n-t_s} \rangle$ and the result follows. □

Next we show that any element of $\mathrm{Syz}(\upsilon_1, \ldots, \upsilon_s)$ is generated by a set of homogeneous syzygies.

LEMMA 2.5.5. *Let $\lambda \in \mathrm{Syz}(v_1, \ldots, v_s)$. Then λ can be expressed as a linear combination of homogeneous syzygies.*

PROOF. We prove the result by induction on s. For $s = 1$, $\mathrm{Syz}(v_1) = \mathrm{ann}_R(v_1) = \langle p^{n-t_1}\mathbf{e}_1\rangle$, and the hypothesis is satisfied. For $s = 2$, $\mathrm{Syz}(v_1, v_2) = \{(\lambda_1, \lambda_2) : \lambda_1 v_1 + \lambda_2 v_2 = 0\}$. If $\lambda_i = \beta_i p^{k_i}$, $v_i = \theta_i p^{t_i}$ for $i = 1, 2$ as before, then $\lambda_1 v_1 = -\lambda_2 v_2$ implies $\beta_1\theta_1 p^{t_1+k_1} = -\beta_2\theta_2 p^{t_2+k_2}$. Unless some $\beta_i = 0$ or $t_i + k_i \geq n$ for each i, we have $t_1 + k_1 = t_2 + k_2$. If $\beta_i = 0$ for some i then λ has only one non-zero component and thus must be homogeneous. If $t_i + k_i \geq n$ for each i then we may write (λ_1, λ_2) as a sum of homgeneous syzygies: $(\lambda_1, \lambda_2) = (\lambda_1, 0) + (0, \lambda_2)$. Thus each element of $\mathrm{Syz}(v_1, v_2)$ is homogeneous.

Now suppose the result holds for $s = l$, so that any element of $\mathrm{Syz}(v_1, \ldots, v_l)$ can be expressed as a linear combination of homogeneous syzygies. Let $\lambda = (\lambda_1, \ldots, \lambda_{l+1}) \in \mathrm{Syz}(v_1, \ldots, v_{l+1})$ with $\lambda_i = \beta_i p^{k_i}$, $v_i = \theta_i p^{t_i}$ and $\beta_i \in R^* \cup \{0\}, \theta_i \in R^*$, for $i = 1, \ldots, l+1$. If $\lambda_{l+1} = 0$, then λ can be viewed as an element of $\mathrm{Syz}(v_1, \ldots, v_l)$ and so is a linear combination of homogeneous syzygies by hypothesis. Otherwise, we show that it is possible to choose a homogeneous syzygy $\lambda^h = (\lambda_1^h, \ldots, \lambda_{l+1}^h)$ such that $\lambda_{l+1} \in \langle \lambda_{l+1}^h\rangle$. Then $\lambda = a\lambda^h + (\lambda_l, 0)$ for some $a \in R, \lambda_l \in \mathrm{Syz}(v_1, \ldots, v_l)$, and so λ is a linear combination of homogeneous syzygies. We can always find an appropriate $\lambda^h \in \mathrm{Syz}(v_1, \ldots, v_{l+1})$ in the form of either $p^{n-t_{l+1}}\mathbf{e}_{l+1}$ or $p^{t_{i,l+1}-t_i}\theta_i^{-1}\mathbf{e}_i - p^{t_{i,l+1}-t_{l+1}}\theta_{l+1}^{-1}\mathbf{e}_{l+1}$, and then we determine λ_l by $(\lambda_l, 0) = \lambda - a\lambda^h$. Indeed, if the only homogeneous syzygy with a non-zero $(l+1)$th coefficient has all other entries zero then we may take $\lambda^h = p^{n-t_{l+1}}\mathbf{e}_{l+1}$. Otherwise, choose $i \in \{1, \ldots, l\}$ such that $t_{i,l+1}$ is minimal. We claim that $p^{k_{l+1}} \in \langle p^{t_{i,l+1}-t_{l+1}}\rangle$, so that $k_{l+1} \geq t_{i,l+1} - t_{l+1}$, and thus $\lambda_{l+1} \in \langle \lambda_{l+1}^h\rangle$. This is trivially true if $t_{i,l+1} = t_{l+1}$. Suppose otherwise, so that $t_{i,l+1} = t_i > t_{l+1}$, and $k_{l+1} + t_{l+1} < t_i$ for $i = 1, \ldots, l$. Then there exist $\tau_i \in R^*$ such that

$$\lambda_1 v_1 + \cdots + \lambda_l v_l = -\lambda_{l+1} v_{l+1} = -\tau_{l+1} p^{k_{l+1}+t_{l+1}} \notin \langle p^{k_i+t_i}\rangle \qquad \text{for each} \quad i \in \{1, \ldots, l\}$$

while on the other hand

$$\lambda_1 v_1 + \cdots + \lambda_l v_l = \tau_1 p^{k_1+t_1} + \cdots + \tau_l p^{k_l+t_l} \in \langle p^{k_i+t_i}\rangle \qquad \text{for each} \quad i \in \{1, \ldots, l\}.$$

This gives a contradiction and the result follows. □

Lemmas 2.5.4 and 2.5.5 have the following consequence.

COROLLARY 2.5.6. *$\mathrm{Syz}(v_1, \ldots, v_s)$ has the homogeneous generating set*

$$\{p^{n-t_i}\mathbf{e}_i : i = 1, \ldots, s\} \cup \{p^{t_{i,j}-t_i}\theta_i^{-1}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\mathbf{e}_j : 1 \leq i < j \leq s\}.$$

DEFINITION 2.5.7. Let $v_i X_i, i = 1, \ldots, s$ be monomials and let $\mathbf{h} = (h_1, \ldots, h_s)$ be contained in $\mathrm{Syz}(v_1 X_1, \ldots, v_s X_s)$. We say that $\mathbf{h}$ is homogeneous of degree X if each non-zero h_i is a monomial and X is a term in $R[\mathbf{x}]$ such that $\mathrm{lt}(h_i)X_i = X$ for all $i = \{1, \ldots, s\}$.

LEMMA 2.5.8. *With notation as above $\mathrm{Syz}(v_1 X_1, \ldots, v_s X_s)$ has a finite generating set of homogeneous syzygies.*

PROOF. Since $R[\mathbf{x}]^s$ is Noetherian, $\mathrm{Syz}(v_1 X_1, \ldots, v_s X_s)$ has a finite generating set. Let $\mathbf{h} \in \mathrm{Syz}(v_1 X_1, \ldots, v_s X_s)$ so that $h_1 v_1 X_1 + \cdots + h_s v_s X_s = 0$. By expanding the polyno-

mials h_i we may collect together all those monomials in the sum which share the same term X. Let the corresponding vector be denoted by $\mathbf{h}^X$. Then

$$h_1^X \upsilon_1 X_1 + \cdots + h_s^X \upsilon_s X_s = 0$$

and the required representation is $\sum_X \mathbf{h}^X$. □

We return now to the proof of Theorem 2.5.2. The argument given here is a direct proof, avoiding the complexity of the saturated subsets approach taken, for example, in Adams and Loustaunau (1994, Section 4.2).

Proof of Theorem 2.5.2. It is clear that

$$\left\langle \{p^{n-t_i}\mathbf{e}_i : i = 1, \ldots, s\} \cup \left\{ p^{t_{i,j}-t_i}\theta_i^{-1}\frac{X_{ij}}{X_i}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\frac{X_{ij}}{X_j}\mathbf{e}_j : 1 \le i < j \le s \right\} \right\rangle$$

is contained in $\mathrm{Syz}(\upsilon_1 X_1, \ldots, \upsilon_s X_s)$. For the converse, let $h = (\rho_1 Y_1, \ldots, \rho_s Y_s)$ be a homogeneous element of $\mathrm{Syz}(\upsilon_1 X_1, \ldots, \upsilon_s X_s)$ for monomials $\rho_i Y_i$ with $X_i Y_i = Y$ for $i = 1, \ldots, s$. Then

$$\upsilon_1\rho_1 X_1 Y_1 + \cdots + \upsilon_s\rho_s X_s Y_s = (\upsilon_1\rho_1 + \cdots + \upsilon_s\rho_s)Y = 0$$

so $(\rho_1, \ldots, \rho_s) \in \mathrm{Syz}\,(\upsilon_1, \ldots, \upsilon_s)$. Now

$$\begin{aligned}\sum_{i=1}^{s} \rho_i \mathbf{e}_i &= \sum_{i=1}^{s} \beta_i p^{n-t_i}\mathbf{e}_i + \sum_{i<j} \delta_{ij}(p^{t_{i,j}-t_i}\theta_i^{-1}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\mathbf{e}_j) \\ &= \sum_{i=1}^{s} \beta_i p^{n-t_i}\mathbf{e}_i + \sum_{i=1}^{s}\sum_{j\neq i} \varepsilon_{ij}\delta_{ij} p^{t_{i,j}-t_i}\theta_i^{-1}\mathbf{e}_i\end{aligned}$$

for some $\beta_i, \delta_{ij} \in R$, where $\varepsilon_{ij} = 1$ for $i < j, \varepsilon_{ij} = -1$ for $i > j$ and $\delta_{ij} = \delta_{ji.}$ But then

$$\begin{aligned}\sum_{i=1}^{s} \rho_i Y_i \mathbf{e}_i &= \sum_{i=1}^{s} \beta_i p^{n-t_i} Y_i \mathbf{e}_i + \sum_{i=1}^{s}\sum_{j\neq i} \varepsilon_{ij}\delta_{ij} p^{t_{i,j}-t_i}\theta_i^{-1} Y_i \mathbf{e}_i \\ &= \sum_{i=1}^{s} \beta_i p^{n-t_i} Y_i \mathbf{e}_i + \sum_{i<j} \delta_{ij}(p^{t_{i,j}-t_i}\theta_i^{-1} Y_i \mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1} Y_j \mathbf{e}_j) \\ &= \sum_{i=1}^{s} \beta_i p^{n-t_i} Y_i \mathbf{e}_i + \sum_{i<j} \delta_{ij}\left(p^{t_{i,j}-t_i}\theta_i^{-1}\frac{X_{ij}}{X_i}\mathbf{e}_i - p^{t_{i,j}-t_j}\theta_j^{-1}\frac{X_{ij}}{X_j}\mathbf{e}_j \right)\frac{Y}{X_{ij}}.\end{aligned}$$

By Lemma 2.5.8, $\mathrm{Syz}(\upsilon_1 X_1, \ldots, \upsilon_s X_s)$ has a finite generating set of homogeneous syzygies so that

$$\{p^{n-t_i}\mathbf{e}_i : i = 1, \ldots, s\} \cup \left\{ p^{t_{i,j}-t_i}\theta_i^{-1}\frac{X_{ij}}{X_i}\mathbf{e}_i - p^{t_{i,j}-tj}\theta_j^{-1}\frac{X_{ij}}{X_j}\mathbf{e}_j : 1 \le i < j \le s \right\}$$

forms a homogeneous generating set for $\mathrm{Syz}(\upsilon_1 X_1, \ldots, \upsilon_s X_s)$, as required. □

Definition 2.5.9. Let f_1, f_2 be non-zero elements of $R[\mathbf{x}]$ with $\mathrm{lm}(f_i) = \theta_i p^{t_i} X_i$, $i = 1, 2$. Then

$$S(f_1, f_2) = p^{t_{1,2}-t_1}\theta_1^{-1}\frac{X_{12}}{X_1}f_1 - p^{t_{1,2}-t_2}\theta_2^{-1}\frac{X_{12}}{X_2}f_2$$

is called the S-polynomial of f_1 and f_2, where $t_{1,2} = \max\{t_1, t_2\}$, and $X_{12} = \mathrm{lcm}\{X_1, X_2\}$.

We give our extension of Buchberger's algorithm in the form of the following theorem, which is a direct result of Theorems 2.5.1 and 2.5.2.

THEOREM 2.5.10. *Let $\mathcal{G} = \{g_i\}_1^s$ be a set of non-zero polynomials in $R[\mathbf{x}]$ with $\mathrm{lm}(g_i) = \theta_i p^{t_i} X_i, i = 1, \ldots, s$. Then $\mathcal{G}$ is a Gröbner basis if and only if*

$$S(g_i, g_j) \rightarrow_{+\mathcal{G}} 0$$
$$p^{n-t_i} g_i \rightarrow_{+\mathcal{G}} 0$$

for all $1 \leq i < j \leq s$.

EXAMPLE 2.5.11. Let $R[\mathbf{x}] = \mathbf{Z}_{27}[x, y]$ and let $<$ denote degree lexicographic term order with $y < x$. Let $\mathcal{G} = \{g_j\}_{j=1}^4 = \{9, x + 1, 3y^2, y^3 + 13y^2 - 12\}$ as in Example 2.4.6. We apply Theorem 2.5.10 to show that $\mathcal{G}$ is a Gröbner basis. The S-polynomial of g_3 and g_4 is given by

$$S(g_3, g_4) = yg_3 - 3g_4 = 15y^2 + 9$$

and is reducible modulo $\mathcal{G} : 15y^2 + 9 \rightarrow_{g_3} 9 \rightarrow_{g_1} 0$. Similarly

$$S(g_2, g_4) = y^3 g_2 - xg_4 = 14xy^2 + y^3 + 12x,$$

and

$$14xy^2 + y^3 + 12x \rightarrow_{g_4} 14xy^2 + 14y^2 + 12x + 12 \rightarrow_{g_2} 12x + 12 \rightarrow_{g_2} 0.$$

Clearly $S(g_1, g_j) \rightarrow_{+g_1} 0, p^{n-k_j} g_j = 0$ where $\nu_p(lc(g_j)) = p^{k_j}$ for $j \in \{1, 2, 3, 4\}$, and

$$S(g_2, g_3) = 3y^2 g_2 - xg_3 = g_3 \rightarrow_{g_3} 0,$$

so we conclude that $\mathcal{G}$ is a Gröbner basis.

ALGORITHM 2.5.12.

INPUT:
$\mathcal{F} = \{f_i\}_{i=1}^l$ a subset of $R[x]$, a term order $<$
OUTPUT:
$\mathcal{G} = \{g_i\}_{i=1}^s$, a Gröbner basis of $\langle f_1, \ldots, f_l \rangle$
INITIALIZATION:
$\mathcal{G} := \mathcal{F}, \mathcal{S} = \{\{f\} : f \in \mathcal{G}\} \cup \{\{g, h\} : g, h \in \mathcal{G}\}$
MAIN PROGRAM:
WHILE $\mathcal{S} \neq \emptyset$ *DO*
    Choose any $\{f\} \in \mathcal{S}$
    Choose any $\{g, h\} \in \mathcal{S}$
    $k := \nu_p(\mathrm{lc}(f))$
    $u := rem_{\mathcal{G}}(p^{n-k} f)$, $v := rem_{\mathcal{G}}(S(g, h))$
    $\mathcal{S} := \mathcal{S} \setminus (\{f\} \cup \{\{g, h\}\})$
    IF $u \neq 0$ THEN
        $\mathcal{S} := \mathcal{S} \cup (\{u\} \cup \{\{a, u\} : a \in \mathcal{G}\})$
    IF $v \neq 0$ THEN
        $\mathcal{S} := \mathcal{S} \cup (\{v\} \cup \{\{a, v\} : a \in \mathcal{G}\})$
    $\mathcal{G} := \mathcal{G} \cup \{u, v\}$

3. Hamming Metric Decoding of an Alternant Code

We summarize results presented in Byrne and Fitzpatrick (preprint), giving an application of the Gröbner basis theory presented in Section 2 toward the decoding of an alternant code over a Galois ring.

3.1. CONSTRUCTION AND MINIMUM DISTANCE OF THE CODE

We assume the reader is familiar with the basic ideas of coding theory (see, for example, MacWilliams and Sloane, 1977; Pless and Huffman, 1998, for a review of the literature). Let $R = \mathrm{GR}(p^{mn}, p^n)$ be a separable extension of a Galois ring T, so $T = \mathrm{GR}(p^{m'n}, p^n)$ where m' divides m. Let $\mathbf{H}$ be the matrix

$$\mathbf{H} = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{N-1} \\ \gamma_0\alpha_0 & \gamma_1\alpha_1 & \cdots & \gamma_{N-1}\alpha_{N-1} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_0\alpha_0^{r-1} & \gamma_1\alpha_1^{r-1} & \cdots & \gamma_{N-1}\alpha_{N-1}^{r-1} \end{bmatrix}$$

where $r \leq N \leq p^m - 1$, $\gamma = (\gamma_0, \ldots, \gamma_{N-1}), \alpha = (\alpha_0, \ldots, \alpha_{N-1}) \in (R^*)^N$, and $\alpha_i - \alpha_j$ is a unit for $i \neq j$. We define the alternant code $C(N, r, \alpha, \gamma, T)$ of length N with symbols in T as the T-submodule

$$C(N, r, \alpha, \gamma, T) = \{\mathbf{c} \in T^N : \mathbf{Hc} = \mathbf{0}\}.$$

Note that the condition $\alpha_i - \alpha_j \in R^*$ for $i \neq j$ is equivalent to the requirement that the set $\{\alpha_j\}_{j=0}^{n-1}$ be contained in a transversal of the cosets of $\langle p \rangle$ in R. We have the following lower bound on the minimum Hamming distance of the code (de Andrade *et al.*, preprint; Norton and Salagean-Mandache, 1999).

THEOREM 3.1.1. *Let $C(N, r, \alpha, \gamma, T)$ be the alternant code defined by the parity check matrix $\mathbf{H}$ as above. Let* $\mathrm{d}_H(N, r, \alpha, \gamma, T)$ denote the minimum Hamming distance of the code $C(N, r, \alpha, \gamma, T)$. Then $\mathrm{d}_H(N, r, \alpha, \gamma, T) > r$.

3.2. DECODING PROCEDURES

Theorem 3.1.1 shows that $C(N, r, \alpha, \gamma, T)$ is a t-error correcting code where $t = \lfloor \frac{r}{2} \rfloor$. In this section we present a decoding algorithm which determines all error patterns of Hamming weight at most t.

Let $\mathbf{v} = \mathbf{c} + \mathbf{e}$ be a received word, where $\mathbf{c} \in C$ and the error vector $\mathbf{e}$ has Hamming weight at most t, and let $\mathbf{s} = \mathbf{Hv} = \mathbf{He}$ be the syndrome vector. Let $\mathcal{J} \subseteq \{0, \ldots, N-1\}$ be the set of indices of non-zero coefficients of $\mathbf{e}$, so that $|\mathcal{J}| \leq t$. The decoding problem seeks initially to determine $\mathcal{J}$, the set of error locations of the error pattern $\mathbf{e}$, and subsequently the error magnitudes e_j for each $j \in \mathcal{J}$. Define the error polynomial $e = \sum_{j \in \mathcal{J}} e_j x^j$ and the syndrome polynomial $S = \sum_{i=0}^{r-1} \sum_{j \in \mathcal{J}} e_j \gamma_j \alpha_j^i x^i$ in the usual way. The error locator polynomial is

$$\Sigma = \prod_{j \in \mathcal{J}} (1 - \alpha_j x),$$

and the error evaluator polynomial is

$$\Omega = \sum_{j \in \mathcal{J}} e_j \gamma_j \prod_{k \in J, k \neq j} (1 - \alpha_k x).$$

These polynomials are related by the well known key equation

$$\Sigma S \equiv \Omega \bmod x^r$$

and the decoding problem is equivalent to solving this congruence subject to certain conditions.

Consider the $R[x]$-submodule $M \subseteq R[x]^2$ consisting of all solutions of the key equation

$$M = \{(a, b) : aS \equiv b \bmod x^r\}.$$

It is easy to see that M is generated by the set $\{(1, S), (0, x^r)\}$. Using the following theorem and the results of the next section we show that the particular solution (Σ, Ω) can be identified in a Gröbner basis of M with respect to a specified term order. Given the generating set $\{(1, S), (0, x^r)\}$ we invoke Algorithm 2.5.12 to compute the required basis.

THEOREM 3.2.1. *Let $R = R_n = \mathrm{GR}(p^{mn}, p^n)$. For each $i = 0, \ldots, n-1$, let $a_i, b_i \in R[x]$, satisfy:*

(i) $\mathrm{lc}(a_i) \in R^*$ *for each* $i \in \{0, \ldots, n-1\}$.
(ii) $\langle a_i, a_j \rangle = R[x]$ *for each* $i, j \in \{0, \ldots, n-1\}$ *such that* $i \neq j$.
(iii) $\langle a_i, b_i \rangle = R[x]$ *for each* $i \in \{0, \ldots, n-1\}$.

Define a, b by

$$a = \prod_{i=0}^{n-1} a_i \qquad \text{and} \qquad b = \sum_{i=0}^{n-1} p^i A_i b_i, \qquad \text{where} \quad A_i = \prod_{j=0, j\neq i}^{n-1} a_j.$$

Then $\langle a, b \rangle$ has a Gröbner basis of the form

$$\{a_1 \cdots a_{n-1}, pa_2 \cdots a_{n-1}, \ldots, p^{n-1}a_{n-2}, p^{n-1}\}.$$

3.3. THE SOLUTION MODULE

The module of all solutions of the key equation

$$M = \{(a, b) : aS \equiv b \bmod x^r\}$$

has the following related structures:

$$M_k = M \cap \langle p^k \rangle$$
$$L_k = M \backslash M_k$$
$$L^* = L_1 = M \backslash M_1.$$

DEFINITION 3.3.1. Let l be an integer. The term order $<_l$ on $R[x]^2$ is defined as follows:

(i) $(x^i, 0) <_l (x^j, 0)$ and $(0, x^i) <_l (0, x^j)$ for $i < j$.
(ii) $(0, x^j) <_l (x^i, 0)$ if and only if $j \leq i + l$.

Explicitly the terms are ordered as

$$(0,1) <_l (0,x) <_l \cdots <_l (0,x^l) <_l (1,0) <_l (0,x^{l+1}) <_l (0,x) <_l \cdots \text{ for } l \geq 0$$

$$(1,0) <_l (x,0) <_l \cdots <_l (x^{-(l+1)},0) <_l (0,1) <_l (x^{-l},0) <_l (0,x) <_l \cdots \text{ for } l<0.$$

An element $(a,b) \in R[x]^2$ has the leading term *on the left* (respectively *on the right*) if $\mathrm{lt}(a,b)$ has the form $(x^i,0)$ (respectively $(0,x^j)$).

We note the general structure of a Gröbner basis of an arbitrary $R[x]$-submodule of $R[x]^2$.

THEOREM 3.3.2. *Let A be an $R[x]$-submodule of $R[x]^2$. Then A has a Gröbner basis of the form*

$$\{(a_0,b_0),\ldots,(a_{n-1},b_{n-1}),(c_0,d_0),\ldots,(c_{n-1},d_{n-1})\}$$

satisfying

(i) for all $i,j \in \{0,\ldots,n-1\}$, $\mathrm{lm}(a_i,b_i) = (p^i x^{\partial a_i},0)$, $\mathrm{lm}(c_j,d_j) = (0,p^j x^{\partial d_j})$.
(ii) $\partial a_i \le \partial a_j$ for $i \ge j$, $\partial d_i \le \partial d_j$ for $i \ge j$.

In what follows we show that the particular solution (Σ,Ω) required in the decoding problem, as formulated previously, is minimal in L^* with respect to the term order $<_{-1}$, defined by

$$(1,0) < (0,1) < (x,0) < (0,x) < \cdots.$$

The next result gives some conditions under which an element is minimal in some L_l. Recall the natural homomorphism, defined in the Introduction:

$$\mu : R \to k_R : f \to f \bmod p$$

which we extend in the obvious way to $R[x]$, and write $\bar{f}$ for μf.

THEOREM 3.3.3. *Let $(a,b) \in M$ satisfy the following for some integer $k \ge 0$:*

(i) $\partial b < \partial a = \partial(\bar{a}) \le t$.
(ii) $\langle p^k \rangle \subseteq \langle a,b \rangle$.

Then (a,b) is minimal in L_{n-k} with respect to the term order $<_{-1}$.

REMARK 3.3.4. The quotient module, given by

$$\begin{aligned} M/M_k &= \{(a,b) + M_k : (a,b) \in M\} \\ &= \{(a,b) + M_k : (a,b) \in L_{n-k}\}, \end{aligned}$$

defines an equivalence relation for each $k \in \{0,\ldots,n\}$. If $\mathcal{G}_k$ is a Gröbner basis for M_k, then $(a,b) + M_k = (a',b') + M_k$ if and only if $\mathrm{Nf}_{\mathcal{G}_k}(a,b) = \mathrm{Nf}_{\mathcal{G}_k}(a',b')$, so that

$$M/M_k = \{\mathrm{Nf}_{\mathcal{G}_k}(a,b) + M_k : (a,b) \in R[x]^2\}.$$

If (a,b) and (a',b') are both minimal in L_{n-k}, then it is not hard to see that $(a,b)+M_k = (a',b') + M_k$.

If an element (a,b) satisfies the conditions of Theorem 3.3.3 then it is the minimal element of some subset L_{n-k} of M, and thus contained, up to equivalence, in a Gröbner basis $\mathcal{G}$ of M. Moreover, if (a,b) is minimal in L_{n-k}, it is certainly minimal in $L_1 = L^*$ and therefore identifiable, up to equivalence, as the minimal regular element of $\mathcal{G}$.

Given an arbitrary vector $\mathbf{v} \in R^N$, write $\mathbf{v} = \mathbf{v}^{\langle 0\rangle} + p\mathbf{v}^{\langle 1\rangle} + \cdots + p^{n-1}\mathbf{v}^{\langle n-1\rangle}$ where

$$v_j = v_j^{\langle i\rangle} p^i \qquad \text{for some} \quad v_j^{\langle i\rangle} \in R^* \cup \{0\}.$$

Note that this representation is not unique since the $v_j^{\langle i\rangle}$ are not necessarily chosen from among a transversal on the cosets of $\langle p\rangle$. For a received word $\mathbf{v} \in R^N$, corresponding to an error vector $\mathbf{e}$, we decompose the syndrome $\mathbf{s}$ in the same way, as follows:

$$\begin{aligned}\mathbf{s} &= \mathbf{v}\mathbf{H}^t = \mathbf{e}\mathbf{H}^t \\ &= \mathbf{e}^{\langle 0\rangle}\mathbf{H}^t + p\mathbf{e}^{\langle 1\rangle}\mathbf{H}^t + \cdots + p^{n-1}\mathbf{e}^{\langle n-1\rangle}\mathbf{H}^t \\ &= \mathbf{s}^{\langle 0\rangle} + p\mathbf{s}^{\langle 1\rangle} + \cdots + p^{n-1}\mathbf{s}^{\langle n-1\rangle}.\end{aligned}$$

THEOREM 3.3.5. *Let Σ and Ω be the error locator and error evaluator polynomials for some error pattern $\mathbf{e}$. Then for each $i \in \{0, \ldots, n-1\}$ there exist polynomials $\Sigma^{\langle i\rangle}$ and $\Omega^{\langle j\rangle} \in R[x]$ satisfying the following:*

(i) $\Sigma = \prod_{j=0}^{n-1} \Sigma^{\langle j\rangle}$.

(ii) $\Omega = \sum_{i=0}^{n-1} p^i \Psi^{\langle i\rangle}\Omega^{\langle i\rangle}$, *where* $\Psi^{\langle i\rangle} = \prod_{j=0, j\neq i}^{n-1} \Sigma^{\langle j\rangle} = \Sigma/\Sigma^{\langle i\rangle}$.

(iii) $\langle \Sigma^{\langle i\rangle}\rangle + \langle \Sigma^{\langle j\rangle}\rangle = R[x]$ *for* $i \neq j$.

(iv) $\langle \Sigma^{\langle i\rangle}\rangle + \langle \Omega^{\langle i\rangle}\rangle = R[x]$ *for each* $i \in \{0, \ldots, n-1\}$.

(v) $\mathrm{lc}(\Sigma^{\langle j\rangle}) \in R^*$ *for each* $j \in \{0, \ldots, n-1\}$.

Moreover, if for each $j \in \{0, \ldots, n-1\}$ we let $\mathcal{J}_i = \{j : e_j^{\langle i\rangle} \neq 0\}$ then suitable polynomials are given by $\Sigma^{\langle i\rangle} = \prod_{j\in\mathcal{J}_i}(1 - \alpha_j x)$ and $\Omega^{\langle i\rangle}$, the unique polynomial of degree less than r such that $S^{\langle i\rangle}\Sigma^{\langle i\rangle} \equiv \Omega^{\langle i\rangle}$ mod x^r.

From Theorem 3.2.1 we deduce that $\langle \Sigma, \Omega\rangle$ has a Gröbner basis of the form

$$\{\Psi^{\langle 0\rangle}, p\Sigma^{\langle 2\rangle}\cdots\Sigma^{\langle n-1\rangle}, p^2\Sigma^{\langle 3\rangle}\cdots\Sigma^{\langle n-1\rangle}, \ldots, p^{n-2}\Sigma^{\langle n-1\rangle}, p^{n-1}\}.$$

In particular, p^{n-1} is contained in $\langle \Sigma, \Omega\rangle$, so the pair Σ, Ω satisfy Condition (ii) of Theorem 3.3.3. We have now proved the following theorem.

THEOREM 3.3.6. *The solution (Σ, Ω) of the key equation required for decoding the alternant code $C(N, r, \alpha, \gamma, S)$ is, up to equivalence, the minimal regular element of a Gröbner basis for the solution module M, under the term order $<_{-1}$.*

Given an element (a, b), minimal in L^*, we compute the roots of Σ as follows. Since $\mu(a, b) = \mu(\Sigma, \Omega)$ for all (a, b) minimal in L^* then in particular

$$\bar{a} = \bar{\Sigma} = \prod_{j\in\mathcal{J}} (1 - \bar{\alpha}_j x).$$

The roots α_j are then determined uniquely from the roots $\bar{\alpha}_j$ and location vector $\alpha = [\alpha_0, \ldots, \alpha_{N-1}]$, whose components comprise a set of distinct coset representatives for the cosets of $\langle p\rangle$ in R.

EXAMPLE 3.3.7. Let R be the Galois ring $Z_4[x]/\langle f\rangle \simeq Z_4[\xi]$ where ξ is a root of the primitive basic irreducible $f = x^4 - 2x^2 + 3x + 1$. Let $C(15, 4, \alpha, \gamma, Z_4)$ be the alternant code whose parity check matrix is defined by $\alpha = [1, \xi, \xi^2, \ldots, \xi^{14}]$ and $\gamma = 1$. Suppose that a codeword $\mathbf{c}$ is sent and the vector $\mathbf{v}$ received. Let $\mathbf{e}$ denote the corresponding error pattern, and suppose that

$$\mathbf{e} = [0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0].$$

The error locator polynomial is given by

$$\Sigma = (1 - \xi^3 x)(1 - \xi^{10} x) = \xi^{13} x^2 + (\xi^{12} + 2\xi^5)x + 1$$

and the syndrome polynomial is

$$S = (\xi^9 + 2)x^3 + (\xi^6 + 2\xi^5)x^2 + (\xi^3 + 2\xi^{10})x + 3.$$

Multiplying Σ by S and reducing modulo x^4 we find that

$$\Omega = (\xi^{10} + 2\xi^{12})x + 3.$$

With the notation of Theorem 3.3.5, $\mathcal{J}_0 = \{3\}, \mathcal{J}_1 = \{10\}, \Sigma^{\langle 0\rangle} = 1 - \xi^3 x$ and $\Sigma^{\langle 1\rangle} = 1 - \xi^{10} x$. If we choose

$$\begin{aligned}\mathbf{e}^{\langle 0\rangle} &= [0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]\\ \mathbf{e}^{\langle 1\rangle} &= [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0]\end{aligned}$$

then the corresponding values for the $S^{\langle i\rangle}$ are given by

$$\begin{aligned}S^{\langle 0\rangle} &= \xi^9 x^3 + \xi^6 x^2 + \xi^3 x + 1\\ S^{\langle 1\rangle} &= x^3 + \xi^5 x^2 + \xi^{10} x + 1\end{aligned}$$

and we get the equations

$$\begin{aligned}\Sigma^{\langle 0\rangle} S^{\langle 0\rangle} &\equiv 1 \mod x^4\\ \Sigma^{\langle 1\rangle} S^{\langle 1\rangle} &\equiv 1 \mod x^4.\end{aligned}$$

Then

$$\begin{aligned}\Psi^{\langle 0\rangle}\Omega^{\langle 0\rangle} + 2\Psi^{\langle 1\rangle}\Omega^{\langle 1\rangle} &= \Sigma^{\langle 1\rangle}\Omega^{\langle 0\rangle} + 2\Sigma^{\langle 0\rangle}\Omega^{\langle 1\rangle}\\ &= 1 - \xi^{10} x + 2(1 - \xi^3 x)\\ &= (\xi^{10} + 2\xi^{12})x + 3\\ &= \Omega.\end{aligned}$$

From Theorem 3.2.1, we deduce that $\{\Sigma^{\langle 1\rangle}, 2\}$ is a Gröbner basis of $\langle \Sigma, \Omega\rangle$ so that by Theorem 3.3.3, the pair (Σ, Ω) is minimal in L^* with respect to the term order $<_{-1}$, and is thus contained, up to equivalence, in a Gröbner basis of M. We now start the decoding proper, applying Theorem 2.5.10 in order to compute the required basis. The module M is generated by $\{(1, S), (0, x^4)\}$. Let $b_1 = (0, x^4)$ and let $b_2 = (1, S)$. Then

$$S(b_1, b_2) = (3x, (\xi^6 + 2\xi^9)x^3 + (\xi^3 + 2\xi^{12})x^2 + x).$$

Applying the division algorithm to $S(b_1, b_2)$ and we find that

$$S(b_1, b_2) \to_{b_2} ((\xi^9 + 2\xi^7)x + \xi^6 + 2\xi^5, 2\xi^7 x^2 + 2\xi^6 x + \xi^6 + 2\xi^9)$$

which is reduced modulo $\{b_1, b_2, b_3\}$. We normalize the remainder, denote it by b_3 and add it to the generating set

$$b_3 = ((\xi^2 + 2\xi^8)x + (\xi^{14} + 2\xi^2), 2x^2 + 2\xi^{14}x + \xi^{14} + 2\xi^{13}).$$

Multiplying by 2 annihilates the leading coefficient, and we obtain

$$b_4 = (2x + 2\xi^{12}, 2\xi^{12})$$

after multiplication by an appropriate unit in R. Now

$$S(b_2, b_3) = ((\xi^2 + 2)x^2 + (\xi^{14} + 2\xi^{13})x + 2\xi^6, 2\xi^5 x^2 + (\xi^{14} + 2\xi^{11})x + 2\xi^6)$$

and

$$S(b_2, b_3) \to_{+\{b_3, b_4\}} (\xi^2 x^2 + (\xi + 2\xi^4)x + \xi^4 + 2\xi^{11}, (\xi^{14} + 2)x + \xi^4 + 2\xi^{13}).$$

For convenience, we multiply by ξ^{13} to get

$$b_5 = (x^2 + (\xi^{14} + 2\xi^2)x + (\xi^2 + 2\xi^9), (\xi^{12} + 2\xi^{14})x + (\xi^2 + 2\xi^{11}))$$

with which we augment the generating set. The S-polynomial

$$S(b_1, b_3) = ((\xi^2 + 2)x^3 + (\xi^{14} + 2\xi^{13})x^2, 2\xi^{14}x^3 + (\xi^{14} + 2\xi^2)x^2)$$

reduces to $(0, 0)$ modulo $\{b_3, b_4, b_5\}$ and

$$S(b_4, b_5) = (2\xi^5 x + 2\xi^2, 2\xi^2) \to_{b_4} (0, 0).$$

Thus the set $\{b_1, b_2, b_3, b_4, b_5\}$ and hence the set $\{b_2, b_3, b_4, b_5\}$ is a Gröbner basis for M. Note that the form of this basis has been predicted in Theorem 3.3.2. The minimal regular element of this basis is b_5, so that $b_5 + M_1 = (\Sigma, \Omega) + M_1$, where $M_1 = M \cap \langle 2 \rangle$. The roots of $x^2 + \bar{\xi}^{12}x + \bar{\xi}^2$ are $\{\bar{\xi}^5, \bar{\xi}^{12}\}$, where $\bar{\xi} = \mu\xi$, and correspond to the error locations $\{3, 10\}$. We implement a modification of Forney's procedure (Forney, 1965; Interlando *et al.*, 1997) to recover the error magnitudes

$$e_i = \frac{S_1 + \Sigma_{i,1} S_0}{\alpha_i + \Sigma_{i,1}}$$

where the coefficients $\Sigma_{i,1}$ are determined by the relation

$$\Sigma_{i,1} = \Sigma_1 + \alpha_i.$$

Then

$$\begin{aligned} e_3 &= \frac{S_1 + \Sigma_{3,1} S_0}{\xi^3 + \Sigma_{3,1}} \\ &= \frac{\xi^3 + 2\xi^{10} - (\xi^{13} + 2\xi^5 + \xi^3)}{\xi^3 + \xi^{12} + 2\xi^5 + \xi^3} \\ &= (\xi^{12} + 2\xi^{11})(\xi^3 + 2\xi^2) = 1 \end{aligned}$$

and

$$\begin{aligned} e_{10} &= \frac{S_1 + \Sigma_{10,1} S_0}{\xi^{10} + \Sigma_{10,1}} \\ &= \frac{\xi^3 + 2\xi^{10} - (\xi^{12} + 2\xi^5 + \xi^{10})}{\xi^{10} + \xi^{12} + 2\xi^5 + \xi^{10}} \\ &= (2\xi^{12})(\xi^3 + 2\xi^6) = 2. \end{aligned}$$

References

Adams, W. W., Loustaunau, P. (1994). An introduction to Gröbner bases. *Graduate Stud. Math.*, **3**.
de Andrade, A., Interlando, J. C., Palazzo, R., Jr. On alternant codes over commutative rings, preprint.
Becker, T., Weispfenning, V. (1993). *Gröbner Bases: A Computational Approach to Commutative Algebra*. New York, Springer-Verlag.
Byrne, E., Fitzpatrick, P. Hamming metric decoding of alternant codes defined over Galois rings, preprint.
Calderbank, A. R., Sloane, N. J. A. (1995). Modular and p-adic cyclic codes. *Des. Codes Cryptogr.*, **6**, 21–35.
Cox, D., Little, J., O'Shea, D. (1992). *Ideals, Varieties, and Algorithms*. New York, Springer-Verlag.
Fitzpatrick, P. (1995). On the key equation. *IEEE Trans. Inf. Theory*, **41**, 1290–1302.
Fitzpatrick, P. (1996). On the scalar rational interpolation problem. *Math. Control Signals Syst.*, **9**, 352–369.
Fitzpatrick, P. (1997). Solving multivariable congruences by change of term order. *J. Symb. Comput*, **24**, 505–510.
Fitzpatrick, P. (1999). Errors and erasures decoding of BCH codes. *IEEE Proc.*, IEE Proc., Commun., **146**, 79–81.
Fitzpatrick, P., Jennings, S. M. (1998). Comparison of two algorithms for decoding alternant codes. *Appl. Algebra Eng. Commun. Comput.*, **9**, 211–220.
Forney, G. D., Jr. (1965). On decoding BCH codes. *IEEE Trans. Inf. Theory*, **11**, 549–557.
Hammons, A. R., Kumar, V., Calderbank, A. R., Sloane, N. J. A., Solé, P. (1994). The Z_4-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Trans. Inf. Theory*, **40**, 301–318.
Interlando, J. C., Palazzo, R., Jr., Elia, M. (1997). On the decoding of Reed–Solomon and BCH codes over integer residue rings. *IEEE Trans. Inf. Theory*, **43**, 1013–1021.
Kanwar, P., Lopez-Permouth, S. R. (1997). Cyclic codes over the integers modulo p^m. *Finite Fields Appl.*, **3**, 334–352.
MacWilliams, F. J., Sloane, N. J. A. (1977). *The Theory of Error-correcting Codes*. Amsterdam, North Holland.
McDonald, B. R. (1974). *Finite Rings with Identity*. New York, Marcel Dekker.
Norton, G. H., Salagean-Mandache, A. (1999). On the key equation over a commutative ring. *Des. Codes Cryptogr.*, **20**, 125–141, to appear.
Pless, V., Huffman, W. (1998). *Handbook of Coding Theory*. Amsterdam, Elsevier.
Raghavendran, R. (1969). Finite associative rings. *Compositio Mathematica*, **21**, 195–229.
Shankar, P. (1979). On BCH codes over arbitrary integer rings. *IEEE Trans. Inf. Theory*, **25**, 480–483.

Originally Received 12 June 2000
Accepted 1 March 2001