Applications of Grobner Bases to Linear Codes.

Jeanne Fitzgerald

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/6248

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPLICATIONS OF GRÖBNER BASES
TO LINEAR CODES

A Dissertation
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Mathematics

by
Jeanne Fitzgerald
B.A., Virginia Polytechnic Institute, 1972
M.S., Virginia Commonwealth University, 1992
August, 1996
Acknowledgments

I am indebted to the Graduate School and to the Department of Mathematics for the opportunity and means to study here at Louisiana State University. In addition, I thank my committee members, Professors William A. Adkins, Richard Kurtz, Richard A. Litherland, O. Carruth McGehee, and James G. Oxley, for all their help.

I am especially grateful to my kind and patient dissertation advisor, Professor Robert F. Lax, with whom it has been my privilege to work.
# Table of Contents

Acknowledgments ................................................................. ii

Abstract .............................................................................. iv

Chapter

1. Affine Variety Codes ......................................................... 1
   1.1 Definition of Affine Variety Codes .................................. 1
   1.2 Viewing Linear Codes as Affine Variety Codes ................. 9

2. Decoding Using Gröbner Bases ........................................... 17
   2.1 Introduction .............................................................. 17
   2.2 Decoding Linear Codes ................................................. 17
   2.3 Precomputation with Syndromes as Variables ................. 33

3. Minimum Distance Calculation ......................................... 51
   3.1 Introduction .............................................................. 51
   3.2 Calculation of $N_0$ .................................................... 52
   3.3 A Lower Bound on Minimum Distance ......................... 68

References ........................................................................... 76

Vita .................................................................................... 79
Abstract

Put $A = \mathbb{F}_q[x_1, \ldots, x_s]$, and let $I$ be an ideal of $A$. Let $P_1, \ldots, P_n$ be all the $\mathbb{F}_q$-rational points of $V(I)$. Define a map $\varphi : A/I \rightarrow \mathbb{F}_q^n$ by $\varphi(\tilde{f}) = (f(P_1), \ldots, f(P_n))$, where $f$ is any preimage of $\tilde{f}$ under the canonical map from $A$ to $A/I$. Let $\{\tilde{f}_i | i \in \mathbb{N}\}$ be a basis of $A/I$ as an $\mathbb{F}_q$-vector space. Define the affine variety code $C$ and its dual $C^\perp$ by

$$ C = \varphi((\tilde{f}_1, \ldots, \tilde{f}_m)); $$

$C^\perp$ is the orthogonal complement of $C$ with respect to the usual inner product in $\mathbb{F}_q^n$.

We show that any linear code can be expressed as an affine variety code.

When a code $C$ is represented as an affine variety code, problems of decoding and finding the minimum distance of $C$ may be expressed as questions about polynomial ideals.

Using the theory of Gröbner bases, along with computer programs that calculate Gröbner bases, we show how to decode and find the minimum distance of any linear code.
Chapter 1
Affine Variety Codes

1.1 Definition of Affine Variety Codes

Let $\mathbf{F}_q$ be the finite field with $q$ elements ($q = p^r$ for some prime $p$ and some positive integer $r$). Let $S \subseteq \mathbf{F}_q[x_1, \ldots, x_s]$. A solution, or zero, of $S$ is an $s$-tuple $(a_1, \ldots, a_s) \in \mathbf{F}_q^s = \mathbf{F}_q \times \mathbf{F}_q \times \cdots \times \mathbf{F}_q$ ($s$ factors), where $\mathbf{F}_q$ is an algebraically closed extension field of $\mathbf{F}_q$ and $f(a_1, \ldots, a_s) = 0$ for all $f$ in $S$. The set of all zeros of $S$, denoted $V(S)$, is called the affine variety in $\mathbf{F}_q^s$ defined by $S$ [Hu, p. 409]. In symbols,

$$V(S) := \{(a_1, \ldots, a_s) \in \mathbf{F}_q^s : f(a_1, \ldots, a_s) = 0 \forall f \in S\}.$$ 

The elements $(a_1, \ldots, a_s)$ of $V(S)$ are known as the points of $V(S)$. Those points of $V(S)$, all of whose coordinates lie in $\mathbf{F}_q$, are called the $\mathbf{F}_q$-rational points of $V(S)$.

Note that if $I$ is the ideal of $\mathbf{F}_q[x_1, \ldots, x_s]$ generated by $S$, then $V(I) = V(S)$. Also note that, in our terminology, a variety need not be an irreducible subset of affine space.

Conversely, for a subset $Y \subseteq \mathbf{F}_q^s$, define $\mathcal{I}(Y)$ by

$$\mathcal{I}(Y) := \{f \in \mathbf{F}_q[x_1, \ldots, x_s] : f(a_1, \ldots, a_s) = 0 \forall (a_1, \ldots, a_s) \in Y\}.$$ 

By Hilbert's Nullstellensatz [Hu, p. 412], $\mathcal{I}(V(I)) = \sqrt{I}$, where $\sqrt{I}$ denotes the radical of $I$. Also, note that $V(I) = V(\mathcal{I}(V(I))) = V(\sqrt{I})$ [Hu, p. 409, Lemma 7.1].

**DEFINITION 1.1.1** Let $I$ be any ideal of $\mathbf{F}_q[x_1, \ldots, x_s]$. Define

$$I_q := I + (x_1^q - x_1, \ldots, x_s^q - x_s).$$
PROPOSITION 1.1.2 Let $I$ be an ideal of $\mathbb{F}_q[x_1, \ldots, x_s]$. Then the points of $V(I_q)$ are the $\mathbb{F}_q$-rational points of $V(I)$.

PROOF. Because $I$ is contained in $I_q$, the variety $V(I_q)$ is contained in the variety $V(I)$. Let $(a_1, \ldots, a_s)$ be a point of $V(I_q)$. Since the polynomial $f_i = x_i^q - x_i$ is an element of $I_q$, $f_i(a_1, \ldots, a_s) = 0$. Substituting $a_i$ for $x_i$, we have $a_i^q - a_i = 0$. A well-known result of finite field theory [LN1, Lemma 2.3, Theorem 2.6, pp. 44-45] states that $a_i \in \mathbb{F}_q$ if and only if $a_i^q - a_i = 0$. Thus $a_i \in \mathbb{F}_q$ for $1 \leq i \leq s$. By definition, $(a_1, \ldots, a_s)$ is an $\mathbb{F}_q$-rational point of $V(I)$.

Conversely, suppose $(a_1, \ldots, a_s)$ is an $\mathbb{F}_q$-rational point of $V(I)$. Then since $a_i \in \mathbb{F}_q$ for $1 \leq i \leq s$, we have $a_i^q - a_i = 0$, by the same result quoted earlier. Thus, $f_i(a_1, \ldots, a_s) = 0$ where $f_i = x_i^q - x_i$. Since $(a_1, \ldots, a_s) \in V(I)$, $g(a_1, \ldots, a_s) = 0$ for all $g \in I$. We can write any polynomial $G(x_1, \ldots, x_s)$ in the ideal $I + (x_1^q - x_1, \ldots, x_s^q - x_s)$ in the form

$$G(x_1, \ldots, x_s) = g(x_1, \ldots, x_s) + \sum_{i=1}^{s} h_i(x_1, \ldots, x_s)(x_i^q - x_i),$$

for some $g(x_1, \ldots, x_s) \in I$ and $h_i(x_1, \ldots, x_s) \in \mathbb{F}_q[x_1, \ldots, x_s]$, $1 \leq i \leq s$. Since $g(a_1, \ldots, a_s) = 0$ and $f_i(a_1, \ldots, a_s) = 0$ (where $f_i = x_i^q - x_i$), it follows that $G(a_1, \ldots, a_s) = 0$, and therefore $(a_1, \ldots, a_s) \in V(I_q)$. \hfill $\Box$

Since $\mathbb{F}_q$ consists of $q$ elements, $V(I)$ contains at most $q^s$ $\mathbb{F}_q$-rational points; hence the number of points of $V(I_q)$ is finite.

DEFINITION 1.1.3 Let $I$ be an ideal of $\mathbb{F}_q[x_1, \ldots, x_s]$. If the number of points of $V(I)$ is finite, then $I$ is said to be a zero-dimensional ideal [AL, pp. 63-64].

REMARK Clearly, $I_q$ is a zero-dimensional ideal.

For the proof of the next result we will need Seidenberg's Lemma 92 [Sei], stated here as given in Becker and Weispfenning [BW, p. 341].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Lemma 1.1.4 (Seidenberg's Lemma 92) Let $J$ be a zero-dimensional ideal of $F_q[x_1, \ldots, x_s]$, and assume that for $1 \leq i \leq s$, $J$ contains a polynomial $f_i \in F_q[x_i]$ with $\gcd(f_i, f'_i) = 1$. Then $J$ is an intersection of finitely many maximal ideals. In particular, $J$ is then radical.

Corollary 1.1.5 Let $I$ be an ideal of $F_q[x_1, \ldots, x_s]$. Then $I_q$ is a radical ideal.

Proof. Since we already know $I_q$ is a zero-dimensional ideal, it suffices to show that $\gcd(f_i, f'_i) = 1$ for $f_i = x_i^q - x_i$, $1 \leq i \leq s$. The formal derivative $f'_i$ of $f_i$ is $qx_i^{q-1} - 1$. Since $q = p^r$ and we are working in characteristic $p$, $f'_i = 0x_i^{q-1} - 1 = -1$, which means $\gcd(f_i, f'_i) = 1$. Therefore, by Lemma 1.1.4, $I_q$ is radical. □

Note that for any point $P = (a_1, \ldots, a_s) \in V(I_q)$, and for any polynomial $f \in F_q[x_1, \ldots, x_s]$, when we calculate $f(P) = f(a_1, \ldots, a_s)$, we obtain an element of $F_q$, because all the coordinates $a_i$ are elements of $F_q$, and so are all the coefficients of $f$.

Let $R = F_q[x_1, \ldots, x_s]/I_q$. The ring $R$ is called the coordinate ring of the variety $V(I_q)$, and the polynomial functions in $R$ are called regular functions on $V(I_q)$. Designate the elements of $R$ as usual by $\tilde{f}$, where $\tilde{f}$ denotes the equivalence class $f + I_q$, for $f \in F_q[x_1, \ldots, x_s]$. Note that $R$ is an $F_q$-vector space.

Lemma 1.1.6 The dimension of $R$ as an $F_q$-vector space is $n$, the number of points in $V(I_q)$.

Proof. Since $I_q$ is a radical ideal and $V(I_q) = \{P_1, \ldots, P_n\}$, we have

$I_q = M_1 \cap M_2 \cap \cdots \cap M_n$.
where $M_i$ is the maximal ideal of $F_q[x_1, \ldots, x_s]$ corresponding to the point $P_i$, $i = 1, \ldots, n$. Put $A = F_q[x_1, \ldots, x_s]$. Then, by the Chinese Remainder Theorem [Hu, Corollary III.2.27, p. 132],

$$R = A/I_q \cong A/M_1 \oplus A/M_2 \oplus \cdots \oplus A/M_n.$$  

If $P_i = (a_{i1}, \ldots, a_{is})$, then $M_i = (x_1 - a_{i1}, \ldots, x_s - a_{is})$. Hence $A/M_i \cong F_q$, since $a_{i1}, \ldots, a_{is}$ are in $F_q$. Thus $\dim R = n$ as an $F_q$-vector space. \hfill \Box

Let $A^n = F^n_q$ be affine $n$-space over $F_q$, where $n$ is the number of points of $V(I_q)$. Fix an ordering $P_1, \ldots, P_n$ of the points $P_i$ of $V(I_q)$.

Define a map $\phi : R \to A^n$ by

$$\phi(f) = (f(P_1), \ldots, f(P_n)),$$

where $\bar{f} = f + I_q$. We have already noted that $f(P_i) \in F_q$ for $f \in F_q[x_1, \ldots, x_s]$ and $P_i \in V(I_q)$. So $\phi(\bar{f}) \in A^n$. Because $R$ is a quotient of vector spaces, we must take care that $\phi$ is well-defined. Suppose $\bar{f}_1 = \bar{f}_2$. Then $f_2 = f_1 + g$, where $g$ is a polynomial in $I_q$. Then, for $P_i \in V(I_q)$,

$$f_2(P_i) = (f_1 + g)(P_i) = f_1(P_i) + g(P_i).$$

Since $g \in I_q$, $g(P_i) = 0$ for any $P_i \in V(I_q)$. So $f_2(P_i) = f_1(P_i) + 0 = f_1(P_i)$. Therefore, $\phi(\bar{f}_1) = \phi(\bar{f}_2)$, and $\phi$ is well-defined.

LEMMA 1.1.7 The map $\phi$ is an isomorphism of $F_q$-vector spaces.

PROOF. We simply use the definitions of polynomial addition and scalar multiplication to see that $\phi$ is linear.

To show that $\phi$ is one-to-one, suppose that $\phi(\bar{f}_1) = \phi(\bar{f}_2)$. Then $\phi(\bar{f}_1 - \bar{f}_2) = \bar{0} \in A^n$; that is, $f_1 - f_2$ is 0 on every point of $V(I_q)$. By the Nullstellensatz, that implies $f_1 - f_2$ is an element of $\sqrt{I_q}$. But $\sqrt{I_q} = I_q$, since $I_q$ is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
a radical ideal [Corollary 1.1.5]. Since \( f_1 - f_2 \in I_q \), \( \bar{f}_1 - \bar{f}_2 = 0 \), which means \( \bar{f}_1 = \bar{f}_2 \).

By Lemma 1.1.6 \( R \) has dimension \( n \) as an \( F_q \)-vector spaces, as does \( A^n \). Since \( \phi \) is one-to-one, the image of \( \phi \) also has dimension \( n \). Consequently, \( \phi \) is onto and is therefore an isomorphism.

Recall that a linear code of length \( n \) over \( F_q \) is a vector subspace of \( F_q^n \).

Now let \( I \) be any ideal of \( F_q[x_1, \ldots, x_s] \). Let \( L \) be an \( F_q \)-vector subspace of \( F_q[x_1, \ldots, x_s]/I_q \), and let \( P_1, \ldots, P_n \) be all the points of \( V(I_q) \).

**Definition 1.1.8** We define the affine variety codes \( C(I, L) \) and \( C^-(I, L) \) as follows:

\[
C(I, L) = \phi(L),
\]

\[
C^-(I, L) = \phi(L)^-,
\]

where \( \phi(L) \) is the image of the subspace \( L \) under the map \( \phi \), and \( \phi(L)^- \) is the orthogonal complement of \( \phi(L) \) in \( A^n \) with respect to the usual inner product on \( A^n \).

**Remark** In general, a different ordering of the points \( P_1, \ldots, P_n \) of the variety \( V(I_q) \) gives a different, but equivalent, linear code.

**Remark** If \( F_0 \) is a subfield of \( F_q \), we may also consider the subfield subcodes obtained by taking the intersections \( C(I, L) \cap F_0^n \) or \( C^-(I, L) \cap F_0^n \).

We will show in the next section that every linear code may be represented as an affine variety code. We end this section with some examples showing that some well-known classes of codes may naturally be viewed as affine variety codes.

**Example 1.1.9** Reed-Solomon codes [LG, p. 19]. Let \( s = 1 \). Take \( I = (0) \).

Then \( I_q = (x^q - x) \) and \( V(I_q) \) is exactly the set of \( F_q \)-rational points of \( \bar{F}_q^1 \).
So \( V(I_q) \) is \( \mathbb{F}_q \), and \( n = q \). Let \( \alpha \) be a generator of \( \mathbb{F}_q^\times \), where \( \mathbb{F}_q^\times \) is the multiplicative group of \( \mathbb{F}_q \). Order the elements of \( V(I_q) \) as follows:

\[
P_1 = \alpha^1, P_2 = \alpha^2, \ldots, P_{q-1} = \alpha^{q-1} = 1, P_q = 0.
\]

Let \( C \) be the linear subspace of \( \mathbb{A}^q = \mathbb{F}_q^q \) consisting of the images

\[
(f(P_1), \ldots, f(P_q))
\]

of the polynomials \( f \) of degree less than \( k \) in \( \mathbb{F}_q[x] \). As a basis for this set of polynomials we may take the set \( \{1, x, x^2, \ldots, x^{k-1}\} \subseteq \mathbb{F}_q[x]/I_q \) and call the span of this set \( L \). Then the affine variety code \( C = C(I, L) \) is equivalent to an extended Reed-Solomon code of dimension \( k \). This definition of Reed-Solomon codes as affine variety codes is compatible with the original approach of Reed and Solomon. They noted that since a univariate polynomial cannot have more than \( k - 1 \) zeros if its degree is less than \( k \), the minimum weight, and hence the minimum distance, of \( C \) is at least \( n - k + 1 \). Note that the dual code \( C^\perp \) of a Reed-Solomon code is also a Reed-Solomon code.

To obtain a non-extended Reed-Solomon code, take \( I = (x^{q-1} - 1) \). Then \( I_q = (x^{q-1} - 1, x^{q} - x) = I \), since \( x^{q} - x = (x^{q-1} - 1)x \). Including the polynomial \( x^{q-1} - 1 \) in \( I \) restricts the points \( P_i \) of \( V(I_q) \) to the nonzero elements of \( \mathbb{F}_q \).

**Example 1.1.10** Generalized Reed-Muller codes [DGM, p. 409]. These codes are analogues of extended Reed-Solomon codes (viewed as in Example 1.1.9) with the variety \( V(I_q) = \mathbb{F}_q^s \) instead of \( \mathbb{F}_q \). So \( I = (0) \) and \( I_q = (x_1^q - x_1, \ldots, x_s^q - x_s) \); the points of \( V(I_q) \) are the points of \( \mathbb{F}_q^s \), and \( L \) is the subspace of \( \mathbb{F}_q[x_1, \ldots, x_s]/I_q \) of all \( s \)-variable polynomials of total degree at most \( v \), for some integer \( v \) with \( 0 \leq v \leq s(q-1) \). Then \( C(I, L) \) is the \( v \)-th order generalized Reed-Muller code \( R(v, s) \) over \( \mathbb{F}_q \).
Example 1.1.11 One-point algebraic-geometric codes. An algebraic-geometric code (AG code), also called a geometric Goppa code, is a linear subspace of $\mathbb{F}_q^n$, specified using ideas of algebraic geometry. We let $X$ be a nonsingular, absolutely irreducible projective curve defined over $\mathbb{F}_q$, with $P_1, \ldots, P_n$ $\mathbb{F}_q$-rational points of $X$, and let $D$ be the divisor $P_1 + P_2 + \cdots + P_n$. Let $G$ be a divisor rational over $\mathbb{F}_q$ with support disjoint from $D$. Define $L(G)$ as usual to be the linear subspace of the function field of $X$ consisting of the zero function along with all functions $f$ satisfying $(f) + G \geq 0$, where $(f)$ is the divisor of $f$. In other words, if we write $G = \sum_{i=1}^r n_iQ_i - \sum_{i=1}^s m_iQ_i'$ with all $n_i, m_i \geq 0$, we include all functions $f$ in the function field which have zeros of order at least $m_i$ at $Q_i'$ and that may have poles of order at most $n_i$ at $Q_i$. $L(G)$ is a finite-dimensional vector space over $\mathbb{F}_q$ [LG, p. 46]. Then [LG, p. 55] the linear code $C(D, G)$ of length $n$ associated to the pair $(D, G)$ is the image of the linear map $\psi : L(G) \rightarrow \mathbb{F}_q^n$ defined by $f \mapsto (f(P_1), \ldots, f(P_n))$.

A one-point AG code is an AG code $C(D, G)$ with $G = mP$ for some point $P$ of $X$ and some integer $m$. The divisor $D$ is $P_1 + \cdots + P_n$, where $\{P_1, \ldots, P_n\} \cup \{P\}$ is the set of all the $\mathbb{F}_q$-rational points of $X$.

By embedding $X$ in a possibly larger-dimensional projective space in a particular way, we will be able to view a one-point AG code $C(D, G)$ as an affine variety code. If $P$ is the only point of $X$ that lies on the hyperplane at infinity, then the rational functions in $L(G)$ are polynomial functions on the affine curve $X_a$, which contains all the points $P_1, \ldots, P_n$ of $D$. We want, therefore, to embed $X$ in a possibly larger-dimensional projective space so that this is the case.

Let $g$ denote the genus of $X$. Let $P$ be an $\mathbb{F}_q$-rational point of $X$. Consider the complete linear system on $X$ defined by the divisor $(2g + 1)P$. By the
Riemann-Roch Theorem [Ha, IV, Section 3], this linear system has projective
dimension $g + 1$ and defines an embedding

$$\psi : X \rightarrow \mathbb{P}^{g+1},$$

(where $\mathbb{P}^{g+1}$ is projective $g + 1$-space over $\mathbb{F}_q$). Moreover, homogeneous coor-
dinates $Y_0, Y_1, \ldots, Y_{g+1}$ can be chosen on this projective space so that

$$\psi(X) \cap \{Y_0 = 0\} = \{\psi(P)\}.$$

Put $X_a = \psi(X) \setminus \{\psi(P)\}$. Then $X_a$ is a curve in $\mathbb{A}^{g+1} = \mathbb{P}^{g+1} \setminus \{Y_0 = 0\}$. Put $I = \mathcal{I}(X_a) \subseteq \mathbb{F}_q[y_1, \ldots, y_{g+1}]$, where $y_i = Y_i/Y_0$ for $i = 1, \ldots, g + 1$.

The functions in the linear space $L(m\psi(P))$ are certain polynomials in the
variables $y_1, \ldots, y_{g+1}$. Saints [Sai, pp. 130-134] showed that the one-point AG
codes $C(D, mP)$ and $C(D', m\psi(P))$ are the same, where $D$ is the sum of the
$\mathbb{F}_q$-rational points of $X \setminus \{P\}$ and $D'$ is the sum of the $\mathbb{F}_q$-rational points of
$\psi(X) \setminus \{\psi(P)\}$. Since the functions in $L(m\psi(P))$ are polynomials, the one-
point AG code $C(D, mP)$ is the affine variety code $C(I, L)$, where $L$ is the
image of $L(m\psi(P))$ in $\mathbb{F}_q[y_1, \ldots, y_{g+1}]/I_q$.

**Remark** Affine variety codes are very closely related to the evaluation codes
considered by Høholdt, van Lint, and Pellikaan [HLP]. Put $A = \mathbb{F}_q[x_1, \ldots, x_s]$ and let $I$ be an ideal of $A$. Let $P_1, \ldots, P_n$ be $\mathbb{F}_q$-rational points of $V(I)$. (In
[HLP], these are not necessarily all the $\mathbb{F}_q$-rational points of $V(I)$). Define a
map $\varphi : A/I \rightarrow \mathbb{F}_q^n$ by $\varphi(\tilde{f}) = (f(P_1), \ldots, f(P_n))$, where $\tilde{f}$ is any preimage
of $\tilde{f}$ under the canonical map from $A$ to $A/I$. Let $\{f_i| i \in \mathbb{N}\}$ be a basis of
$A/I$ as an $\mathbb{F}_q$-vector space. Define the evaluation code $E_m$ and its dual $C_m$ by

$$E_m = \varphi(\tilde{f}_1, \ldots, \tilde{f}_m)$$

$$C_m = E^L_m.$$
Let \( f_1, \ldots, f_m \) be linearly independent elements in \( R = A/I_q \). Let \( \tilde{f}_i, i = 1, \ldots, m, \) be a preimage in \( A/I \) of \( f_i \) under the canonical map from \( A/I \) to \( A/I_q \). Extend \( \tilde{f}_1, \ldots, \tilde{f}_m \) to a basis of \( A/I \). Put \( L = (\langle \tilde{f}_1, \ldots, \tilde{f}_m \rangle) \). Then it is easy to see that the AV code \( C(I, L) \) (respectively, \( C^L(I, L) \)) is equal to the evaluation code \( E_m \) (respectively \( C_m \)), where \( P_1, \ldots, P_n \) are all the \( \mathbb{F}_q \)-rational points of \( V(I) \).

1.2 Viewing Linear Codes as Affine Variety Codes

In this section we show that any linear code can be viewed as an affine variety code. We already showed, in section 1 of this chapter, that if there are exactly \( n \) \( \mathbb{F}_q \)-rational points \( P_1, \ldots, P_n \) in the affine variety \( V(I) \subseteq \mathbb{F}_q^s \), (so that \( |V(I)| = n \)), then there is an \( \mathbb{F}_q \)-vector space isomorphism

\[
\phi : \mathbb{F}_q[x_1, \ldots, x_s]/I_q \longrightarrow \mathbb{A}^n
\]

given by

\[
\phi(\tilde{f}) = (f(P_1), \ldots, f(P_n)).
\]

Because \( \phi \) is onto, for any \( n \)-tuple \( (a_1, \ldots, a_n) \) whose coordinates are elements of \( \mathbb{F}_q \), there is an element \( \tilde{f} \in \mathbb{F}_q[x_1, \ldots, x_s]/I_q \) such that

\[
f(P_1) = a_1, f(P_2) = a_2, \ldots, f(P_n) = a_n.
\]

Assuming that we know a variety \( V(I) \) with exactly \( n \) \( \mathbb{F}_q \)-rational points, we may view each of the \( m \) rows of any \( m \times n \) matrix whose entries come from \( \mathbb{F}_q \) as the image under the map \( \phi \) of some element of \( \mathbb{F}_q[x_1, \ldots, x_s]/I_q \). In particular, we may express parity check matrices and generator matrices of linear codes in this way.

We intend, then, to construct a variety with exactly \( n \) \( \mathbb{F}_q \)-rational points. In the process, we will also obtain explicit formulas for the desired elements \( \tilde{f} \) of \( \mathbb{F}_q[x_1, \ldots, x_s]/I_q \).
First, we need some notation and a few lemmas.

Let $J$ designate the ideal generated by the zero polynomial in $\mathbb{F}_q[x_1, \ldots, x_s]$. As we have already observed, $V = V(J_q)$ is the set of points of $\mathbb{F}_q^s$. There are exactly $q^s$ such points. By Lemma 1.1.7, $\mathbb{F}_q[x_1, \ldots, x_s]/J_q \cong \mathbb{F}_q^s$ by the $\mathbb{F}_q$-vector space isomorphism $\phi$, where $\phi(f)$ is the evaluation of $f$ on the points of $\mathbb{F}_q^s$.

**Definition 1.2.1** Let $P_1, \ldots, P_{q^s}$ be the points of $\mathbb{F}_q^s$. The characteristic function

$$e_{P_i} : V \rightarrow \mathbb{F}_q$$

is the function that sends $P_i$ to $1 \in \mathbb{F}_q$ and $P_j$ to $0 \in \mathbb{F}_q, i \neq j$.

Clearly, any function $f : V \rightarrow \mathbb{F}_q$ can be written as an $\mathbb{F}_q$-linear combination of characteristic functions.

**Lemma 1.2.2** [DGM, p 406]. Suppose $P_i = (a_{i1}, a_{i2}, \ldots, a_{is})$, with $a_{ik} \in \mathbb{F}_q$ for $1 \leq k \leq s$. Define a polynomial $E_{P_i} \in \mathbb{F}_q[x_1, \ldots, x_s]$ by $E_{P_i} = \prod_{k=1}^{s}(1 - (x_k - a_{ik})^{q-1})$. Then $E_{P_i}$ restricted to $\mathbb{F}_q^s$ is the characteristic function $e_{P_i}$.

**Proof.** Since $a^{q-1} = 1$ for any nonzero $a \in \mathbb{F}_q$, the polynomial $1 - (x_k - a_{ik})^{q-1} = 0$ whenever $x_k \neq a_{ik}$; thus the polynomial $E_{P_i}$ is the same as the characteristic function $e_{P_i}$ on the points of $\mathbb{F}_q^s$.

The characteristic functions $e_{P_i}$, $i = 1, \ldots, q^s$, span the $\mathbb{F}_q$-vector space $\mathbb{F}_q[x_1, \ldots, x_s]/J_q = R$, and since there are $q^s$ of them, they form an $\mathbb{F}_q$-vector space basis for $R$.

Now let $W$ be a subset of $V$ and suppose $|W| = n \leq q^s$. By renumbering the $P_i$ if necessary, we may assume that $W = \{P_1, \ldots, P_n\}$.

As an $\mathbb{F}_q$-vector space $R = \mathbb{F}_q[x_1, \ldots, x_s]/J_q = V_1 \oplus V_2$ where $V_1$ is the vector subspace generated by $\{e_{P_1}, \ldots, e_{P_n}\}$ and $V_2$ is the vector subspace generated by $\{e_{P_{n+1}}, \ldots, e_{P_{q^s}}\}$. Note that the characteristic functions generating
$V_2$ are zero on all the points $P_1, \ldots, P_n$ of $W$. That means the corresponding $E_{P_{n+1}}, \ldots, E_{P_s}$ are also zero on all the points of $W$, and therefore they are elements of $\mathcal{I}(W)$.

**Lemma 1.2.3** Let $I$ be the ideal $(E_{P_{n+1}}, \ldots, E_{P_s})$. Then $I_q = I + (x_1^q - x_1, \ldots, x_s^q - x_s) = \mathcal{I}(W)$.

**Proof.** We noted in the paragraph above that $E_{P_{n+1}}, \ldots, E_{P_s}$ are all elements of $\mathcal{I}(W)$, and certainly the polynomials $x_1^q - x_1, \ldots, x_s^q - x_s$ are also elements of $\mathcal{I}(W)$, because every point of $W$ is an $\mathbb{F}_q$-rational point. Therefore $I_q \subseteq \mathcal{I}(W)$.

To show that $\mathcal{I}(W) \subseteq I_q$, we prove the contrapositive. Suppose a polynomial $f$ is not an element of $I_q$. Consider $\bar{f}$ in $R = \mathbb{F}_q[x_1, \ldots, x_s]/J_q$. Since the functions $e_{P_i}$, $i = 1, \ldots, q^s$, are a basis for $R$, $\bar{f} = \sum_{i=1}^{q^s} a_i e_{P_i}$, where the $a_i$ are elements of $\mathbb{F}_q$. So $f = (\sum_{i=1}^{q^s} a_i E_{P_i}) + g$, where $g$ is some element of $J_q$. Thus, $f = \sum_{i=1}^{q^s} a_i E_{P_i} + \sum_{j=n+1}^{q^s} b_j E_{P_j} + g$. Since $\sum_{j=n+1}^{q^s} b_j E_{P_j} + g$ is an element of $I_q$, there must be some $a_i$, $1 \leq i \leq n$, with $a_i \neq 0$, because we have assumed $f$ is not an element of $I_q$. Say $a_{i_0} \neq 0$. Then $f$ is nonzero on $P_{i_0} \in W$, and consequently, $f$ cannot be an element of $\mathcal{I}(W)$. $\square$

We have constructed an ideal $I_q$ such that $V(I_q)$ has exactly $n$ $\mathbb{F}_q$-rational points, as we wanted. Note that since $I_q = \mathcal{I}(W)$, $I_q$ is a radical ideal. The following theorem summarizes the results of the discussion above.

**Theorem 1.2.4** For any set $\{P_1, \ldots, P_n\}$ in $\mathbb{F}_q^s$, we can construct a generating set $S$ for a radical ideal $I_q$ of $\mathbb{F}_q[x_1, \ldots, x_s]$ such that $V(S) = V(I_q) = \{P_1, \ldots, P_n\}$. Moreover, given any $n$-tuple $(b_1, \ldots, b_n)$ of elements of $\mathbb{F}_q$, we can construct a polynomial $f$ in $\mathbb{F}_q[x_1, \ldots, x_s]$ such that $f(P_i) = b_i$.

**Proof.** We have already proved the first statement in the theorem. To
construct the polynomial \( f \), take \( f = \sum_{i=1}^{n} b_i E_i \). \( \square \)

**Definition 1.2.5** Let \( W = \{P_1, \ldots, P_n\} \) be an affine variety \( \subseteq \mathbb{F}_q^* \). Let \( f_1, \ldots, f_m \) be polynomials in \( \mathbb{F}_q[x_1, \ldots, x_s] \). We call the matrix \( M = [a_{ij}] \), where \( a_{ij} = f_i(P_j) \), a polynomial evaluation matrix for \( W \).

**Theorem 1.2.6** We can construct any \( m \times n \) matrix \( M \), whose entries come from some finite field \( \mathbb{F}_q \), as a polynomial evaluation matrix for some affine variety \( W \) contained in \( \mathbb{F}_q^* \) for some \( s \), with \( |W| = n \).

**Proof.** Take \( s \) to be a positive integer such that \( n \leq q^s \). Let \( P_1, \ldots, P_n \) be the points of \( \mathbb{F}_q^* \). Take \( W \) to be the points \( P_1, \ldots, P_n \); by our previous results, \( W = V(S) \), where \( S = \{x_1^q - x_1, \ldots, x_s^q - x_s, E_{P_{s+1}}, \ldots, E_{P_{s'}}\} \).

Suppose \( M = [a_{ij}] \), \( a_{ij} \in \mathbb{F}_q \). Let \( f_i, 1 \leq i \leq m \), be the polynomial given by \( \sum_{j=1}^{n} a_{ij} E_j \). Then \( f_i(P_j) = a_{ij} \), and \( M \) is a polynomial evaluation matrix on \( W \). \( \square \)

**Example 1.2.7** Constructing an identity matrix using polynomials.

Let \( W \) be the \( \mathbb{F}_2 \)-rational points of \( \mathbb{F}_2^3 \). Note that

\[
W = \{P_1 = (0,0,0), P_2 = (0,0,1), P_3 = (0,1,0), \ldots, P_8 = (1,1,1)\}.
\]

In this case, as we have already seen more generally, \( \mathcal{I}(W) \) is the radical ideal \( (x_1^2 - x_1, x_2^2 - x_2, x_3^2 - x_3) \) of \( \mathbb{F}_2[x_1, x_2, x_3] \).

We next calculate the polynomials \( E_{P_i} \), \( i = 1, \ldots, 8 \). For example, to calculate \( E_{P_2} \), we note that \( P_2 = (0,0,1) \); so by Lemma 1.2.3,

\[
E_{P_2} = (1 - (x_1 - 0)^1)(1 - (x_2 - 0)^1)(1 - (x_3 - 1)^1) = x_1 x_2 x_3 + x_1 x_3 + x_2 x_3 + x_3.
\]

Consider the \( q^s \times q^s = 8 \times 8 \) matrix \( M \) with \( M = [E_i(P_j)] \), \( i = 1, \ldots, 8 \), \( j = 1, \ldots, 8 \). This matrix, shown below, is the \( 8 \times 8 \) identity matrix.
Theorem 1.2.8 Any linear code is an affine variety code.

Proof. Any linear code \( C \) is a vector subspace of some finite-dimensional \( \mathbb{F}_q \)-vector space, so we can find a parity check matrix \( M \) for \( C \). By Theorem 1.2.6, we can represent \( M \) as a polynomial evaluation matrix \([f_i(P_j)]\) for some finite affine variety \( W = \{P_1, ..., P_s\} \) contained in \( \mathbb{F}_q^s \), for some integer \( s \), and polynomials \( f_1, ..., f_m \) of \( \mathbb{F}_q[x_1, ..., x_s] \). Thus, \( C \) is the affine variety code \( C^\perp(I, L) \) where \( I_q = \mathcal{I}(W) \) and \( L \) is the linear subspace of \( \mathbb{F}_q[x_1, ..., x_s]/I_q \) generated by the images of the polynomials \( f_i, i = 1, ..., m \).

Remark. Rather than constructing a parity check matrix of a linear code, we could instead construct its generator matrix as a polynomial evaluation matrix, say with the polynomials \( g_1, ..., g_m \). Letting \( L \) be the linear subspace of \( \mathbb{F}_q[x_1, ..., x_s]/I_q \) generated by the images of the polynomials \( g_1, ..., g_m \), we see that \( C \) is the affine variety code \( C(I, L) \).

Example 1.2.9 Ternary Golay code [Go] A parity check matrix for the ternary Golay code is:

\[
M = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 2 & 1 & 0 & 2 \\
0 & 0 & 1 & 0 & 0 & 1 & 2 & 1 & 0 & 1 & 2 \\
0 & 0 & 0 & 1 & 0 & 1 & 2 & 0 & 1 & 2 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 2 & 2 & 1 & 1
\end{pmatrix}
\]

Of course, the entries in this matrix are elements of \( \mathbb{F}_3 \).
The number of columns of $M$ is $n = 11$. We need an integer $s$ with $n \leq 3^s$, so we take $s = 3$. Rather than use the variables $x_1, x_2,$ and $x_3$, we will use $x, y,$ and $z$.

We order the points of $\mathbb{F}_3^3$ as follows:

\[ P_1 = (0,0,0), P_2 = (0,0,1), P_3 = (0,0,2), P_4 = (0,1,0), \ldots, P_{27} = (2,2,2). \]

The first eleven points are $P_1 = (0,0,0), \ldots, P_{11} = (1,0,1)$.

The ideal $I \subseteq \mathbb{F}_3[x,y,z]$ such that $V(I_3) = \{P_1, \ldots, P_{11}\}$ is generated by the polynomials $E_{P_1}, \ldots, E_{P_{27}}$. We construct each $E_{P_i}$ using the formula in Lemma 1.2.2. For example,

\[ E_{P_1} = E_{(1,0,2)} = (1-(x-1)^2)(1-(y-0)^2)(1-(z-2)^2) = (2x^2+2x)(1+2y^2)(2z^2+z). \]

If we want, we can use software to find a smaller generating set for $I$, or in this case, inspection and some experimentation reveal that we may take $I$ to be the ideal $(xy, x + 2x^2, xz^2 + 2xz)$.

Next, we construct the polynomials $E_{P_1}, \ldots, E_{P_{11}}$, then calculate the polynomial $f_i$ giving the $i$-th row of the matrix $M$ by

\[ f_i = \sum_{j=1}^{11} a_{ij} E_{P_j}, \]

where $a_{ij}$ is the $i,j$ entry of the matrix $M$ above. We then have the following polynomial evaluation matrix $[f_i(P_j)]$ with

\[
\begin{align*}
  f_1 &= 1 + x + y - xz + y^2 - z^2 + y^2z, \\
  f_2 &= y - z + xz - y^2 - yz - z^2 + y^2z, \\
  f_3 &= x - y + z + xz + y^2 + yz - z^2 - y^2z + yz^2 - y^2z^2, \\
  f_4 &= -x + y - xz + y^2z - yz^2 - y^2z^2, \text{ and} \\
  f_5 &= x + yz^2;
\end{align*}
\]
By taking the dimension $s$ of the affine space $\mathbb{F}_q^s$ large enough, we can ensure that the polynomials $f_1, \ldots, f_m$ generating $L$ are monomials.

**Theorem 1.2.10** Any linear code $C$ can be represented as an affine variety code $C^\perp(I, L)$ for some ideal $I \subseteq \mathbb{F}_q[x_1, \ldots, x_s]$ and for some integer $s$, with $L = (f_1, \ldots, f_m)$ and $f_1, \ldots, f_m$ monomials.

**Proof.** Let the code $C$ have $m \times n$ parity check matrix $M = [a_{ij}]$, with $a_{ij} \in \mathbb{F}_q$ for all $i, j$. If all the columns of $M$ are distinct, take $s = m$ (the number of rows of $M$), and for $j = 1, \ldots, n$, take the point $P_j$ to be the $m$-tuple $(a_{1j}, \ldots, a_{mj})$. That is, $P_j$ is the $j$-th column of $M$. Take $P_{n+1}, \ldots, P_{qm}$ to be all the other points of $\mathbb{F}_q^m$, ordered arbitrarily. Take $I = (E_{P_{n+1}}, \ldots, E_{P_{qm}})$, as in the construction of $I$ in Theorem 1.2.6. Take $f_1 = x_1, \ldots, f_m = x_m$, which are certainly monomials.

If the columns of $M$ are not distinct, we must modify this construction. Such a parity check matrix does not bode well for the error-correcting capabilities of our code, but to present the proof in full generality we forge ahead. In this case, we construct a new matrix $M'$ by adding rows to $M$ until we obtain a matrix having distinct columns. For example, if

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

in which columns 1 and 2 are equal, we could construct $M'$ using a row in
which every entry is zero except in the second column, where we use a 1:

\[ M' = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 & 5 \\ 2 & 2 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \]

We now have \( n \) distinct columns, and we proceed as in the first paragraph of this proof. However, we still take only the first \( m \) monomials \( f_1 = x_1, \ldots, f_m = x_m \) to generate \( L \). □
Chapter 2
Decoding Using Gröbner Bases

2.1 Introduction

In this chapter we describe an algorithm for decoding any linear code via computation of a Gröbner basis for a polynomial ideal associated to the code and the desired number of errors to be corrected. Using this algorithm, we can decode linear codes up to the true minimum distance. This decoding method is a generalization of a method proposed by X. Chen, I.S. Reed, T. Helleseth, and T.K. Truong [CRHT1] for decoding binary cyclic codes. In a subsequent paper [CRHT2], they extended their results to cyclic codes over $F_q$ for arbitrary $q$. They also described how to decode (one-point) algebraic-geometric codes from affine plane curves by finding separate error locators for each coordinate of the errors.

2.2 Decoding Linear Codes

As we showed in Chapter 1, we can view a linear code $C$ as the affine variety code $C(I, L)$, where $L = \langle \bar{f}_1, \ldots, \bar{f}_m \rangle$ and $V(I_q) = \{P_1, \ldots, P_n\} \subseteq F_q^n$. Let $I_q$ be generated by $\{g_1, \ldots, g_l\}$, with $g_1, \ldots, g_l \in F_q[x_1, \ldots, x_s]$. We specifically include the polynomials $x_1^q - x_1, \ldots, x_s^q - x_s$ with which we construct $I_q$ from $I$ among the generators $g_1, \ldots, g_l$ of $I_q$. The matrix $M = [f_i(P_j)]$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is a parity check matrix for $C$.

In this context we recall some terms and results from coding theory.

**Definition 2.2.1** Let $r = (r_1, \ldots, r_n) \in F_q^n$. The *syndrome* of $r$ is the
$m$-tuple $S = (S_1, \ldots, S_m)$, where

$$S_i = \sum_{j=1}^{n} r_j f_i(P_j).$$

**Remark 2.2.2** $Mr^T = S^T$. Note that $r$ is a codeword of $C$ if and only if the syndrome of $r$ is $(0, \ldots, 0)$.

**Definition 2.2.3** Let $r = (r_1, \ldots, r_n) \in \mathbb{F}_q^n$. The weight of $r$, denoted $\text{wt}(r)$, is the number of nonzero coordinates of $r$.

Now suppose the codeword $c = (c_1, \ldots, c_n)$ is transmitted over a noisy channel and the received word $r = (r_1, \ldots, r_n) = (c_1, \ldots, c_n) + (e_1, \ldots, e_n)$ is received. To decode $r$, we must determine the error vector $e = (e_1, \ldots, e_n)$. If $\text{wt}(e) = t$, we say that $t$ errors have occurred.

Assume now that precisely $t$ errors have occurred. Also, assume that there exists a unique vector $e = (e_1, \ldots, e_n)$ with $\text{wt}(e) \leq t$, such that $r - e \in C$. This is always true if $t \leq \lfloor \frac{d_{\text{min}} - 1}{2} \rfloor$, where $d_{\text{min}}$ is the minimum distance of $C$.

Let $S = (S_1, \ldots, S_m)$ be the syndrome of $r$. Note that $S$ is also the syndrome of $e$, since

$$S^T = Mr^T = Mc^T + Me^T = (0, \ldots, 0)^T + Me^T,$$

since $c \in C$.

**Lemma 2.2.4** The vector $e$ is the unique element of $\mathbb{F}_q^n$ with $\text{wt}(e) \leq t$ such that $Me^T = S$.

**Proof.** Let $e'$ be an element of $\mathbb{F}_q^n$, with $\text{wt}(e') \leq t$, such that $Me'^T = S$. Then $M(r - e')^T = S - S = (0, \ldots, 0)^T$, so $r - e'$ is a codeword. Since we have assumed that $e$ is the unique vector $x$ of weight less than or equal to $t$ for which $r - x$ is a codeword, $e' = e$. □

We can determine the error vector $e = (e_1, \ldots, e_n)$ if we can determine the $t$ nonzero coordinates of $e$. We need to find both the error locations, that is,
the indices $i$ where $e_i$ is nonzero, and the error values, that is, the elements of $\mathbb{F}_q$ appearing in each of the error locations. We denote the error values by $E_j$, $j = 1, \ldots, t$, and we associate each error location to its corresponding error point of $V(I_q)$. For example, an error in the 5-th coordinate of a received word is associated to the point $P_5$ of the variety $V(I_q)$. The reason for this association is that the summands in the equations $\sum_{j=1}^{n} e_j f_i(P_j) = S_i$ giving the syndromes are nonzero only at these error points. We want to construct a set of polynomials whose zeros tell us the error values and locations. We denote the $t$ error points by the $t$-tuples of indeterminates $(x_{i1}, \ldots, x_{is}), \ldots, (x_{t1}, \ldots, x_{ts})$. We use the $t$ indeterminates $a_1, \ldots, a_t$ for the error values.

**Definition 2.2.5** For a fixed number of errors $t$ and for a nonzero syndrome $S = (S_1, \ldots, S_m)$, define the error locator ideal

$$E = E(t, S) \subseteq \mathbb{F}_q[x_{i1}, \ldots, x_{is}, a_1, \ldots, x_{t1}, \ldots, x_{ts}, a_t]$$

by

$$E = E(t, S) := \left( \sum_{j=1}^{t} a_j f_i(x_{j1}, \ldots, x_{js}) - S_i, a_j^{q-1} - 1, g_h(x_{j1}, \ldots, x_{js}) \right),$$

where $h = 1, \ldots, l$; $i = 1, \ldots, m$; and $j = 1, \ldots, t$.

**Lemma 2.2.6** Let $V(E) \subseteq \mathbb{F}_{q^t}^t$ be the variety of $E$. Let

$$(B_{i1}, \ldots, B_{is}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) \in V(E).$$

Then each $s$-tuple $(B_{j1}, \ldots, B_{js}), 1 \leq j \leq t$, is one of the points $P_1, \ldots, P_n$ of $V(I_q)$.

**Proof.** Each $s$-tuple $(B_{j1}, \ldots, B_{js})$ satisfies all the polynomials $g_1, \ldots, g_t$ that generate $I_q$. \qed
Since each $s$-tuple $(B_{j1}, \ldots, B_{js})$ is an element of $\{P_1, \ldots, P_n\}$, we may write $(B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t)$ as $(P_{k_1}, E_1, \ldots, P_{k_t}, E_t)$ where $P_{k_j} \in \{P_1, \ldots, P_n\}$ for $1 \leq j \leq t$.

**Lemma 2.2.7** The ideal $\mathcal{E}$ is a zero-dimensional, radical ideal of

$$\mathbb{F}_q[x_{i1}, \ldots, x_{is}, a_1, \ldots, x_{t1}, \ldots, x_{ts}, a_t].$$

**Proof.** Each point in $V(\mathcal{E})$ may be written $(P_{k_1}, A_1, \ldots, P_{k_t}, A_t)$, where the $P_{k_j}, 1 \leq j \leq t$, are elements of $\{P_1, \ldots, P_n\}$. Moreover, each $A_j, 1 \leq j \leq t$, must satisfy $A_j^q - 1$, so there are a finite number $(q - 1)$ possibilities for each $A_j$. Thus, there are at most $(q - 1)^t$ elements in $V(\mathcal{E})$. Since $V(\mathcal{E})$ is finite, $\mathcal{E}$ is zero-dimensional by definition. To show that $\mathcal{E}$ is a radical ideal, note that since $A_j^q - 1$ is an element of $\mathcal{E}$ for $j = 1, \ldots, t$, the polynomials $a_j^q - a_j = a_j(a_j^{q-1} - 1)$ are also in $\mathcal{E}$. Also, for each $x_{ji}, i = 1, \ldots, s$, $j = 1, \ldots, t$, the polynomial $x_{ji}^q - x_{ji}$ is an element of $\mathcal{E}$, because we have chosen the generators $g_1, \ldots, g_l$ of $I_q$ to include these polynomials. Then $\mathcal{E}$ meets the conditions of Seidenberg's Lemma 92 (Lemma 1.1.4), so it is radical.

**Proposition 2.2.8** Let $u = (P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \in V(\mathcal{E})$. Then the points $P_{k_1}, \ldots, P_{k_t}$ are distinct elements of $\{P_1, \ldots, P_n\} = V(I_q)$.

**Proof.** Since $u \in V(\mathcal{E})$,

$$\sum_{j=1}^{t} E_j f_i(P_{k_j}) = S_i,$$

for $i = 1, \ldots, m$. If $P_{k_{j_1}} = P_{k_{j_2}}$, we can rewrite this, using the distributive property in $\mathbb{F}_q$. To make the notation simpler, suppose $P_{k_1} = P_{k_2}$, and suppose all the other $P_{k_j}$ are distinct. The following argument easily generalizes (but
with complicated notation) to more difficult cases. For each \( i, 1 \leq i \leq m, \)
\[
\sum_{j=1}^{t_i} E_j f_i(P_{k_j}) = E_1 f_i(P_{k_1}) + E_2 f_i(P_{k_2}) + \sum_{j=3}^{t_i} E_j f_i(P_{k_j}) \\
= (E_1 + E_2) f_i(P_{k_1}) + \sum_{j=3}^{t_i} E_j f_i(P_{k_j}).
\]
Now consider the element \( y = (y_1, \ldots, y_n) \) of \( \text{F}_q^n \), where
\[
y_{k_1} = E_1 + E_2, \\
y_{k_j} = E_j, j = 3, \ldots, t,
\]
and all the other coordinates of \( y \) are zero. Then \( y \) is a vector of weight strictly less than \( t \). (The weight of \( y \) is \( t - 1 \) if \( E_1 + E_2 \neq 0 \) and \( t - 2 \) if \( E_1 + E_2 = 0 \).)
Since the weight of \( y \) is less than \( t, y \neq e \), which has weight exactly \( t \). By leaving out the summands where \( y_j = 0 \), we have
\[
\sum_{j=1}^{n} y_j f_i(P_{j}) = \sum_{j=1}^{t} E_j f_i(P_{k_j}) = S_i.
\]
Thus \( My^T = S^T \). But by Lemma 2.2.4, \( e \) is the unique vector of weight less than or equal to \( t \) such that \( Me^T = S^T \). Since \( y \neq e \), we have reached a contradiction. Therefore, \( P_{k_1} \neq P_{k_2} \).

**Theorem 2.2.9** Let \( u = (P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \in V(\mathcal{E}). \) Let \( \hat{e} = (e_1, \ldots, e_n) \)
where \( e_{k_j} = E_j \) for \( j = 1, \ldots, t \) and all the other coordinates of \( \hat{e} \) are zero.
Then \( \hat{e} = \hat{e}(u) = e, \) the error vector.

**Proof.** Note that by the previous proposition, the weight of \( \hat{e} \) is \( t \) and that, leaving out the summands where \( e_j = 0 \) and rearranging the terms, we have
\[
\sum_{j=1}^{n} e_j f_i(P_{j}) = \sum_{j=1}^{t} E_j f_i(P_{k_j}) = S_i,
\]
so \( Me^T = S^T \). By Lemma 2.2.4, \( \hat{e} = e. \) 

**Remark 2.2.10** Thus, we can decode \( r \) if we can find any element \( u \) of \( V(\mathcal{E}). \)

**Corollary 2.2.11** If \( u = (P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \) and \( u' = (P'_{k_1}, E'_1, \ldots, P'_{k_t}, E'_t) \)
are elements of \( V(\mathcal{E}), \) then there is a permutation \( \sigma \) in \( \text{Sym}(t), \) the symmetric
group on \( t \) elements, such that \( E'_j = E_{\sigma(j)} \) and \( P'_{k_j} = P_{\sigma(j)} \), for \( j = 1, \ldots, t \). Moreover, any \( \sigma \in \text{Sym}(t) \), applied in this way to \( u \in V(\mathcal{E}) \), gives an element \( u' \) of \( V(\mathcal{E}) \).

**Proof.** Let \( u \) and \( u' \) be elements of \( V(\mathcal{E}) \). By Theorem 2.2.9, \( \hat{e}(u) = e = \hat{e}(u') \). The first statement of the corollary follows from the definition of \( \hat{e} \). To prove the second statement, note that, because of the symmetry of \( \mathcal{E} \) with respect to the index \( j \), any permutation \( \sigma \) applied to the index \( j \) leaves the polynomial \( \sum_{j=1}^{t} E_j f_i(P_{k_j}) - s_i \) unchanged and the sets of polynomials \( \{E_j^{q-1} - 1\} \) and \( \{g_h(x_{j_1}, \ldots, x_{j_s})\} \), \( (h = 1, \ldots, l; i = 1, \ldots, m; j = 1, \ldots, t) \) unchanged. \( \square \)

**Definition 2.2.12** Define the projection map \( \pi : F_q^{t+st} \rightarrow F_q^{t+s} \) by

\[
\pi(u) = \pi(X_{11}, \ldots, X_{1s}, A_1, \ldots, X_{ts}, A_t) = (X_{11}, \ldots, X_{1s}, A_1).
\]

If \( u \in V(\mathcal{E}) \), we may rewrite this as

\[
\pi(u) = \pi(P_{k_1}, E_1, \ldots, P_{k_t}, E_t) = (P_{k_1}, E_1).
\]

**Proposition 2.2.13** \( (P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \in V(\mathcal{E}) \) if and only if \( (P_{k_j}, E_j) \in \pi(V(\mathcal{E})) \) for all \( j, 1 \leq j \leq t \).

**Proof.** \((\Rightarrow)\) Suppose \( (P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \in V(\mathcal{E}) \). Clearly, \( (P_{k_1}, E_1) \in \pi(V(\mathcal{E})) \). By Corollary 2.2.11, for any permutation \( \sigma \) in \( \text{Sym}(t) \),

\[
(P_{k_{\sigma(1)}}, E_{\sigma(1)}, \ldots, P_{k_{\sigma(t)}}, E_{\sigma(t)}) \in V(\mathcal{E}).
\]

Let us denote this element, by an abuse of notation, as \( \sigma(u) \). Now let \( \sigma \) be the permutation \((12 \cdots t)\). Then \( \pi(\sigma^{-1}(u)) = (E_j, P_{k_j}) \) for \( j = 1, \ldots, t \), and so \( (P_{k_j}, E_j) \in \pi(V(\mathcal{E})) \) for \( 1 \leq j \leq t \).

\((\Leftarrow)\) Now suppose \( \pi(V(\mathcal{E})) \) is a nonempty set of \( v \) distinct elements \( (P_{k_1}, E_1), \ldots, (P_{k_v}, E_v) \). We first show that \( v = t \). Since \( \pi(V(\mathcal{E})) \) is nonempty, there is
an element \( u \) in \( V(\mathcal{E}) \) with \( u = (P_{u_1}, A_1, \ldots, P_{u_t}, A_t) \), with the points \( P_{u_j}, 1 \leq j \leq t \), distinct (by Proposition 2.3.8). By the forward direction of the proof of this proposition, the projections \( \pi(\sigma(u)) \) are in \( \pi(V(\mathcal{E})) \) for every \( \sigma \) in \( \text{Sym}(t) \), so \((P_{u_1}, A_1), \ldots, (P_{u_t}, A_t)\) are distinct elements of \( \pi(V(\mathcal{E})) \). Therefore \( v \geq t \). But since (by Corollary 2.3.11) every element of \( V(\mathcal{E}) \) is \( \sigma(u) \) for some \( \sigma \) in \( \text{Sym}(t) \), \( v = t \). Hence \( u' = (P_{k_1}, E_1, \ldots, P_{k_v}, E_v) \in V(\mathcal{E}) \). □

Remark 2.2.14 Thus, to decode \( r \), we need only find the elements of \( \pi(V(\mathcal{E})) \). That is, we need to find only the solutions to the unknowns \( a_1 \) and \( x_{11}, \ldots, x_{1s} \).

Put \( I_\pi = \mathcal{E} \cap \mathbb{F}_q[x_{11}, \ldots, x_{1s}, e_1] \).

Lemma 2.2.15 \( I_\pi \) is a radical ideal of \( \mathbb{F}_q[x_{11}, \ldots, x_{1s}, e_1] \).

Proof. Since \( I_\pi \) contains the polynomials \( e_1^{e_1^n-1} - 1 \) and \( x_{11}^q - x_{11}, \ldots, x_{1s}^q - x_{1s} \), \( I_\pi \) is a zero-dimensional ideal. By Seidenberg's Lemma, using the polynomial \( e_1^q - e_1 = e_1(e_1^{e_1^n-1} - 1) \) and the polynomials \( x_{11}^q - x_{11}, \ldots, x_{1s}^q - x_{1s} \), \( I_\pi \) is radical. □

Lemma 2.2.16 [CLO, Lemma 1, p. 121 and Theorem 3, p. 192] Using the notation above, \( \pi(V(\mathcal{E})) = V(I_\pi) \).

Proof. We first show that \( \pi(V(\mathcal{E})) \subseteq V(I_\pi) \). Let \( f \in I_\pi \). If

\[
(P_{k_1}, E_1, \ldots, P_{k_t}, E_t) \in V(\mathcal{E}),
\]

then \( f \) vanishes at \((P_{k_1}, E_1, \ldots, P_{k_t}, E_t)\), since \( f \in \mathcal{E} \). But since \( f \) involves only the variables \( x_{11}, \ldots, x_{1s}, e_1 \), we may write

\[
f(P_{k_1}, E_1) = f(\pi((P_{k_1}, E_1, \ldots, P_{k_t}, E_t))) = 0.
\]

Thus \( f \) vanishes at all points of \( \pi(V(\mathcal{E})) \), and hence \( \pi(V(\mathcal{E})) \subseteq V(I_\pi) \).

For the reverse inclusion, we will show that if \( f \in I(\pi(V(\mathcal{E}))) \), then \( f \in I_\pi \). That means that \( I(\pi(V(\mathcal{E}))) \subseteq I_\pi \), so \( V(I_\pi) \subseteq V(I(\pi(V(\mathcal{E}))) \). Suppose
\[ f = f(x_{11}, \ldots, x_{1s}, e_1) \in I(\pi(V(\mathcal{E}))), \text{ i.e. } \text{f is zero on all points of } \pi(V(\mathcal{E})). \]

Then, considered as an element of \( F_q[x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{t1}, \ldots, x_{ts}, e_t] \), we have \( f(B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) = 0 \) for all \((B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) \in V(\mathcal{E}). \) So \( f \in I(V(\mathcal{E})). \) By Hilbert's Nullstellensatz, \( f^N \in \mathcal{E} \) for some integer \( N. \) But since \( f \) depends only on \( x_{11}, \ldots, x_{1s}, e_1, \) so does \( f^N, \) and so \( f^N \in \mathcal{E} \cap F_q[x_{11}, \ldots, x_{1s}, e_1] = I_\pi. \) Therefore, \( f \in \sqrt{I_\pi}, \) but by Lemma 2.2.15, \( I_\pi \) is radical, so \( f \in I_\pi. \)

We now have \( V(I_\pi) \subseteq V(I(\pi(V(\mathcal{E}))), \) so it remains to show only that \( \pi(V(\mathcal{E})) = V(I(\pi(V(\mathcal{E}))), \) of \( \pi(V(\mathcal{E})) \) is the smallest affine variety containing \( \pi(V(\mathcal{E})), \) we need only demonstrate that \( \pi(V(\mathcal{E})) \) is an affine variety. But since \( \pi(V(\mathcal{E})) \) is a finite set of points, we are done. \( \Box \)

**Corollary 2.2.17** For any intersection ideal \( J_\pi = \mathcal{E} \cap F_q[X], \) where \( X \) is any subset of the set \( \{x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{t1}, \ldots, x_{ts}, e_t\}, \)

1. \( J_\pi \) is a radical ideal.

2. If \( \pi \) is the projection map on the coordinates of \( V(\mathcal{E}) \) corresponding to the variables in the subset \( X, \) then \( \pi(V(\mathcal{E})) = V(J_\pi). \)

**Proof.** Similar to the proofs of Lemmas 2.2.15 and 2.2.16. \( \Box \)

We plan to show that we can find the points of \( \pi(V(\mathcal{E})) \) by computing a Gröbner basis for the ideal \( I_\pi. \) To do this, we now need some definitions and results from the theory of Gröbner bases. For a more complete exposition, see [AL].

Denote by \( T^m \) the set of (monic) monomials in the variables \( x_1, \ldots, x_m. \)

That is,

\[ T^m = \{ x_1^{\beta_1} \cdots x_m^{\beta_m} : \beta_i \in \mathbb{N}, i = 1, \ldots, m \}. \]
To make this notation less cumbersome, we will sometimes write $x^\beta$ for $x_1^{\beta_1} \cdots x_m^{\beta_m}$, where $\beta = (\beta_1, \ldots, \beta_m) \in \mathbb{N}^m$.

**Definition 2.2.18 [AL, pp. 18-19]** A term order on $T^m$ is a total order $<$ on $T^m$ satisfying:

1. $1 < x^\beta$ for all $x^\beta \in T^m$, $x^\beta \neq 1$;

2. If $x^\alpha < x^\beta$, then $x^\alpha x^\gamma < x^\beta x^\gamma$ for all $x^\gamma \in T^m$.

We will say that $<$ is a term order on the variables $x_1, \ldots, x_m$ if $<$ is a term order on $T^m$.

An important example of a term order is the lexicographic order:

**Definition 2.2.19 [AL, p. 19]** The lexicographic order on $T^m$ with $x_1 < x_2 < \cdots < x_m$ is defined as follows. For $\alpha = (\alpha_1, \ldots, \alpha_m)$, $\beta = (\beta_1, \ldots, \beta_m) \in \mathbb{N}^n$, define $x^\alpha < x^\beta$ if and only if the first coordinates $\alpha_i$ and $\beta_i$ in $\alpha$ and $\beta$ from the right, which are different, satisfy $\alpha_i < \beta_i$.

We will also use elimination orders:

**Definition 2.2.20 [AL, p. 69]** Let $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_m\}$ be two sets of variables, and let $<_x$ and $<_y$ be term orders on the monomials in the $x$ variables and the $y$ variables respectively. Then a term order $<$ may be defined on the monomials in $x_1, \ldots, x_n, y_1, \ldots, y_m$ as follows. For $X_1, X_2$ monomials in $x_1, \ldots, x_n$ and $Y_1, Y_2$ monomials in $y_1, \ldots, y_m$, define $X_1Y_1 < X_2Y_2$ if and only if $(X_1 <_x X_2)$ or $(X_1 = X_2$ and $Y_1 <_y Y_2)$. This term order is called an elimination order with the $x$ variables larger than the $y$ variables.

**Remark 2.2.21** Lexicographic order on $\mathbb{F}_q[x_1, \ldots, x_m]$ with $x_1 < x_2 < \cdots < x_m$ is an elimination order with the variables $x_{i+1}, \ldots, x_m$ larger than the variables $x_1, \ldots, x_i$ for any $i$, $1 < i < m$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Suppose we have a term order on $T^m$. Then for all $f \in F_q[x_1, \ldots, x_m]$, with $f \neq 0$, we may write

$$f = a_1x^{a_1} + a_2x^{a_2} + \cdots + a_rx^{a_r},$$

with $0 \neq a_i \in F_q$, $x^{a_1} \in T^m$, and $x^{a_1} > x^{a_2} > \cdots > x^{a_r}$. Define the leading monomial of $f$, denoted $\text{lm}(f)$, by

$$\text{lm}(f) = x^{a_1}.$$

**Definition 2.2.22** [AL, p. 32] A set of nonzero polynomials $G = \{G_1, \ldots, G_t\}$ contained in an ideal $I$ of $F_q[x_1, \ldots, x_m]$ is called a Gröbner basis for $I$ if and only if for all $f \in I$ such that $f \neq 0$, there exists $i \in \{1, \ldots, t\}$ such that $\text{lm}(G_i)$ divides $\text{lm}(f)$.

**Definition 2.2.23** [AL, p 48] A Gröbner basis $G = \{G_1, \ldots, G_t\}$ is called a reduced Gröbner basis if, for all $i$, the coefficient of the leading monomial $\text{lm}(G_i)$ is 1, and, for all $i$, no nonzero term in $G_i$ is divisible by $\text{lm}(G_j)$ for any $j \neq i$.

We accept as given the existence of Gröbner bases as well as algorithms to compute them. The interested reader may consult [AL, pp 25-51]. It is important to note that if $G = \{G_1, \ldots, G_t\}$ is a Gröbner basis for an ideal $I$, then $I = (G_1, \ldots, G_t)$ [AL, Corollary 1.6.3, p. 33]. Note also that since a Gröbner basis $G$ is defined in terms of leading monomials, $G$ depends on the term order chosen for the monomials of $F_q[x_1, \ldots, x_m]$.

The following important theorem allows us to obtain a Gröbner basis for the ideal $I_\pi$.

**Theorem 2.2.24** Let $I$ be a nonzero ideal of $F_q[x_1, \ldots, x_m, y_1, \ldots, y_n]$, and let $\prec$ be an elimination order with the $y$ variables larger than the $x$ variables. Let
$G = \{G_1, \ldots, G_f\}$ be a Gröbner basis for this ideal. Then $G \cap F_q[x_1, \ldots, x_m]$ is a Gröbner basis for the ideal $I \cap F_q[x_1, \ldots, x_m]$.

**Proof.** See [AL, p. 69].

Using Theorem 2.2.24, by choosing an elimination order $<$ on $F_q[x_{i_1}, \ldots, x_{i_s}, e_i]$, with the variables $x_{i_1}, \ldots, x_{i_s}, e_i$ less than all the other variables, and by computing a Gröbner basis for the ideal $\mathcal{E}$ using this elimination order, we may obtain a Gröbner basis for $I_\pi$. Suppose $H = \{H_1, \ldots, H_r\}$ is the Gröbner basis for $I_\pi$, obtained in that way.

Recall that, by definition, the elimination order consists of a term order $<_1$ on the variables $x_{i_1}, \ldots, x_{i_s}, e_i$ and another term order $<_2$ on all the other variables, along with another condition making the first set of variables less than all the other variables. The term orders $<_1$ and $<_2$ may be chosen arbitrarily. Therefore, we can choose the term order $<_1$ for the variables $x_{i_1}, \ldots, x_{i_s}, e_i$ to be the lexicographic term order. For $<_2$ we could use a more computationally efficient order, such as graded reverse lex, cf. [Bu, GMNRT].

The Gröbner basis $H$ we obtain by the method of Theorem 2.2.24 is then a Gröbner basis for $I_\pi$, computed with respect to the lexicographic order. In addition, $H \cap F_q[X]$ is a Gröbner basis for the ideal $\mathcal{E} \cap F_q[X]$ for any subset $X$ of $\{x_{i_1}, \ldots, x_{i_s}, e_i\}$.

**Theorem 2.2.25** Let $<_1$ be the lexicographic term order on the variables $x_{i_1}, \ldots, x_{i_s}, e_i$, with $x_{i_1} < x_{i_2} < \cdots < x_{i_s} < e_i$, and let $<_2$ be any term order on the variables $x_{i_1}, \ldots, x_{i_s}, e_i, e_\pi$. Let $<$ be an elimination order, defined as in Definition 2.2.20, with the variables $x_{i_1}, \ldots, x_{i_s}, e_i$ less than all the other variables. Let $G$ be a Gröbner basis for $\mathcal{E}$ with respect
to $<$. Then we can solve for the points of $\pi(V(\mathcal{E})) = \mathcal{E} \cap \mathbf{F}_q[x_{11}, \ldots, x_{1s}, e_1]$ by applying elimination theory to the Gröbner basis $G$.

**Proof.** Consider the ideal

$$I_{x_{11}} = \mathcal{E} \cap \mathbf{F}_q[x_{11}] = I_\pi \cap \mathbf{F}_q[x_{11}].$$

By Corollary 2.2.17, the variety $V(I_{x_{11}}) \subseteq \mathbf{F}_q$ is precisely the set of first coordinates $B_{11}$ of points $(B_{11}, \ldots, B_{1s}, E_1)$ of $\pi(V(\mathcal{E}))$. Note that $H = \{H_1, \ldots, H_r\}$ is a Gröbner basis for $I_\pi$ under the term order $<$, which is the same as the term order $<_1$ when restricted to the variables $x_{11}, \ldots, x_{1s}, e_1$. By Remark 2.2.21, the term order $<_1$, which is lexicographic order on $x_{11}, \ldots, x_{1s}, e_1$ with $x_{11} < x_{12} < \cdots < x_{1s} < e_1$, is an elimination order with $x_{12}, \ldots, x_{1s}, e_1$ larger than $x_{11}$. Therefore, by Theorem 2.2.24, $H \cap \mathbf{F}_q[x_{11}]$ is a Gröbner basis for $I_{x_{11}}$. Every polynomial in $H \cap \mathbf{F}_q[x_{11}]$ is a univariate polynomial in the variable $x_{11}$. Since $I_{x_{11}}$ is an ideal of $\mathbf{F}_q[x_{11}]$, which is a principal ideal domain, $I_{x_{11}} = (L)$, where $L$ is the monic polynomial of lowest degree in $I_{x_{11}}$. Note that $L$ is an element of $H \cap \mathbf{F}_q[x_{11}]$. This is true because by definition of a Gröbner basis, there is a polynomial $L'$ whose leading monomial divides the leading monomial of $L$. (We may assume that the polynomials of $H$ are monic.) If $L \neq L'$, then the remainder upon division of $L$ by $L'$ is also an element of $I_{x_{11}}$, a contradiction, because we have assumed that $L$ is the monic polynomial of lowest degree in $I_{x_{11}}$. Thus, we may find the set of first coordinates $B_{11}$ of points $(B_{11}, \ldots, B_{1s}, E_1)$ of $\pi(V(\mathcal{E}))$ by finding the roots of the polynomial of lowest degree in $H \cap \mathbf{F}_q[x_{11}]$. We can accomplish this by a Chien search [Ch; LN2, pp. 493-494] or any other method for finding the roots of a univariate polynomial. Let the set of roots of $L$ be $\{R_1, \ldots, R_a\}$. Note that these roots are elements of $\mathbf{F}_q$, for they all satisfy the polynomial $x_{11}^a - x_{11} \in \mathcal{E}$.
Next consider the ideal

$$I_{x_1, x_2} = \mathcal{E} \cap \mathbb{F}_q[x_{11}, x_{12}] = I_x \cap \mathbb{F}_q[x_{11}, x_{12}].$$

By Corollary 2.2.17, $V(I_{x_{11}, x_{12}})$ is precisely the set of pairs $(B_{11}, B_{12})$ of first and second coordinates of points $(B_{11}, \ldots, B_{1s}, E_i)$ of $\pi(V(\mathcal{E}))$. As before, note that $H \cap \mathbb{F}_q[x_{11}, x_{12}]$ is a Gröbner basis for $I_{x_{11}, x_{12}}$, because the lexicographic term order $<_1$ on the variables $x_{11}, x_{12}, \ldots, x_{1s}, e_1$ is an elimination order with the variables $x_{13}, \ldots, x_{1s}, e_1$ larger than the variables $x_{11}$ and $x_{12}$. Every element in $H \cap \mathbb{F}_q[x_{11}, x_{12}]$ is a polynomial in the two variables $x_{11}$ and $x_{12}$. Of course, some are univariate in $x_{11}$. In fact we already know all the common roots $\{R_1, \ldots, R_u\}$ of the univariate polynomials in $x_{11}$. Substituting $R_1$ for the variable $x_{11}$ in all the elements of $H \cap \mathbb{F}_q[x_{11}, x_{12}]$, we obtain a set of univariate polynomials in the variable $x_{12}$. We may, as before, find all the common roots $R_{1,1}, \ldots, R_{1,u}$ of these polynomials. (The substitution of $R_1$ in the univariate polynomials in $x_{11}$ of course yields the polynomial 0, whose roots include every element of $\mathbb{F}_q$.) Substituting the other first coordinates $R_2, \ldots, R_u$ for $x_{11}$ into the elements of $H \cap \mathbb{F}_q[x_{11}, x_{12}]$ in turn and solving for $x_{12}$, we obtain all the first and second coordinates $(R_i, R_{i,j})$ of the points $(B_{11}, \ldots, B_{1s}, E_i)$ of $\pi(V(\mathcal{E}))$.

We continue this process recursively, obtaining univariate polynomials in the variable $x_{1k}$ by substituting in turn every already-determined $(k - 1)$-tuple $(R_i, R_{i,j}, \ldots, R_{i,j}, \ldots, w)$ of first through $(k - 1)$-st coordinates in place of the corresponding variables $x_{11}, x_{12}, \ldots, x_{1(k-1)}$ in all the polynomials of $H \cap \mathbb{F}_q[x_{11}, x_{12}, \ldots, x_{1(k-1)}, x_{1k}]$. Then we find the common roots of all these univariate polynomials in $x_{1k}$ to get the $k$-th coordinates of the points of $\pi(V(\mathcal{E}))$ whose first $(k - 1)$ coordinates are $(R_i, R_{i,j}, \ldots, R_{i,j}, \ldots, w)$.
Although we have given the last coordinate $E_1$ of the points $(B_{11}, \ldots, B_{1s}, E_1)$ a distinctive letter name because it is the "error value" rather than a coordinate of the "error point", the procedure for finding its values is no different from finding the values of any other coordinate of $(B_{11}, \ldots, B_{1s}, E_1)$.

The process stops after we have completed the procedure for all $s + 1$ coordinates of the points of $\pi(V(\mathcal{E}))$. □

The actual solution procedure is easier and more natural to do than to describe. Here is an example.

**Example 2.2.26** Let $C$ be the affine variety code $C^4(I, L)$, where $I = (y^2 + y - x^3) \subseteq \mathbb{F}_4[x,y]$ and $L = (1, x, y, \bar{x}, \bar{y})$. The points of $V(I_4)$ are the $\mathbb{F}_4$-rational points of the curve $y^2 + y - x^3$; they are the eight points $(0, 0), (0, 1), (1, \alpha), (1, \alpha^2), (\alpha, \alpha), (\alpha, \alpha^2), (\alpha^2, \alpha), (\alpha^2, \alpha^2)$, where $\alpha^2 = \alpha + 1$ is a generator of $\mathbb{F}_4^*$. This code $C$ is the same as the geometric Goppa code

$$C_n(P_1 + \cdots + P_8, 5P_\infty),$$

and it has minimum distance 5 [YK]. Therefore $C$ can correct 2 errors.

A parity check matrix for $C$ is

$$M = \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & \alpha & \alpha & \alpha^2 & \alpha^2 \\
0 & 1 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\
0 & 0 & 1 & 1 & \alpha^2 & \alpha & \alpha & \alpha \\
0 & 0 & \alpha & \alpha^2 & \alpha^2 & 1 & 1 & \alpha
\end{pmatrix}$$

Suppose we receive the error vector $(0, 0, 1, 0, \alpha, 0, 0)$ of weight 2. The syndrome of this error vector is $(\alpha^2, \alpha, \alpha^2, 0, 0)$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The error locator ideal \( E_{(a^2, a, a^2, 0, 0)} \) is the ideal generated by the following polynomials:

\[
\begin{align*}
& e_1 + e_2 - a^2, \\
& e_1 x_1 + e_2 x_2 - a, \\
& e_1 y_1 + e_2 y_2 - a^2, \\
& e_1 x_1^2 + e_2 x_2^2, \\
& e_1 x_1 y_1 + e_2 x_2 y_2, \\
& e_1^2 - 1, \\
& e_2^2 - 1, \\
& x_1^3 - x_1, \\
& x_2^3 - x_2, \\
& y_1^3 - y_1, \\
& y_2^3 - y_2, \\
& y_1^3 + y_1 - x_1^3, \\
& y_2^3 + y_2 - x_2^3.
\end{align*}
\]

The lexicographic term order with \( x_1 < y_1 < e_1 < x_2 < y_2 < e_2 \) is an elimination order in all the senses we require. Using the computer program MAS [Kr], we find that a Gröbner basis \( G \) for \( E \), with respect to the lexicographic order with \( x_1 < y_1 < e_1 < x_2 < y_2 < e_2 \) is

\[
\begin{align*}
& x_1^2 + (\alpha + 1)x_1 + \alpha, \\
& y_1 + \alpha x_1, \\
& e_1 + x_1, \\
& x_2 + x_1 + (\alpha + 1), \\
& y_2 + \alpha x_1 + 1, \\
& e_2 + x_1 + (\alpha + 1).
\end{align*}
\]

By Remark 2.2.14, we need consider only the first three polynomials, since they are the only ones in \( G \cap \mathbb{F}_4[x_1, y_1, e_1] \). The polynomial \( x_1^2 + (\alpha + 1)x_1 + \alpha \) is the only element of \( G \cap \mathbb{F}_4[x_1] \). It is univariate in \( x_1 \), of course. Its roots in \( \mathbb{F}_4 \) are 1 and \( \alpha \). Thus, the set of first coordinates of the error points consists of the elements 1 and \( \alpha \).

The polynomials in \( G \cap \mathbb{F}_4[x_1, y_1] \) are \( x_1^2 + (\alpha + 1)x_1 + \alpha \) and \( y_1 + \alpha x_1 \). We first replace \( x_1 \) by 1, which is one of the values for \( x_1 \) that we have previously found. The polynomial \( x_1^2 + (\alpha + 1)x_1 + \alpha \) vanishes (we already know that 1 is one of its roots), and the polynomial \( y_1 + \alpha x_1 \) becomes \( y_1 + \alpha \). Hence, when
$x_1 = 1$, we have $y_1 = \alpha$. Similarly, substituting the other $x_1$-value $\alpha$, we see that $y_1 = \alpha^2$. We have now found the two error points $(1, \alpha)$ and $(\alpha, \alpha^2)$. To complete the decoding, we consider the polynomials in $G \cap \mathbb{F}_4[x_1, y_1, e_i]$, and for the indeterminates $x_1$ and $y_1$ we substitute first 1 and $\alpha$ respectively. We obtain only one nonzero polynomial, the polynomial $e_1 + 1$. So the error value $e_1$ associated with the error point $(1, \alpha)$ is 1. Next we substitute $\alpha$ and $\alpha^2$ for $x_1$ and $y_1$ respectively, obtaining the polynomial $e_1 + \alpha$. Hence, the error value associated with the error point $(\alpha, \alpha^2)$ is $\alpha$. Since $(1, \alpha)$ is the third point in our chosen point order for the code $C$ and $(\alpha, \alpha^2)$ is the sixth point, we know that the error vector has 1 in the third coordinate, $\alpha$ in the sixth coordinate, and zeros elsewhere. Checking this result with the original error vector that produced the syndrome $(\alpha^2, \alpha, 0, 0)$, we see that we have been successful.

Remark 2.2.27 In [CRHT2], the authors suggest decoding algebraic geometry codes, such as the one in Example 2.2.26, by finding the roots $\beta_{i1}, \ldots, \beta_{iu}$ of each of the monic generator polynomials of the ideals $I_{x_i} = E \cap \mathbb{F}_4[x_i]$, $i = 1, \ldots, s$. The error points are then “suitable combinations" $(\beta_{ij1}, \ldots, \beta_{iju})$, where $\beta_{ij}$ is one of the roots of the monic generator of $I_{x_i}$. However, it is easy to construct examples in which the number of possible combinations of the zeros of these polynomials is quite large. In fact, even in the example above, the number of combinations of the values for $x_1$ and $y_1$ is 4. We must choose 2 of them to be error points; there are $\binom{4}{2} = 6$ ways to do that, making it difficult to select the correct error vector.

We must note that we have assumed that the weight of the error vector is exactly $t$. Let us now consider what would happen, using this method, when the weight of the error vector is either greater than or less than $t$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
First suppose that the weight of the actual error vector is \( u > t \), where the code \( C \) can correct \( u \) or more errors. Since the syndrome \( (s_1, \ldots, s_m) \) belongs to an error vector whose weight is greater than \( t \), the system of polynomial equations that generate \( \mathcal{E} \) has no solution. When the Gröbner basis of the ideal \( \mathcal{E} \) is calculated, its polynomials will have no common roots. If the Gröbner basis we have calculated is a reduced Gröbner basis, it will be the polynomial 1 [AL, p. 63, Theorem 2.2.6]. In either case we will be able to detect this situation.

It is tempting, therefore, to try to set up a single system of polynomials that will give us the error points and error values for any number of errors \( v \leq t \). We would expect, in that case, for certain syndromes, that some of the error values would be zero. But as soon as one of the \( E_i \) is allowed to be zero, its associated \( P_k \) can legitimately be any point of \( V \). Thus the procedure of Theorem 2.2.24 would be useless, giving us as error points all the points of \( V \). So we must first assume that \( t = 1 \). Then if there are no elements in \( \pi(V(\mathcal{E})) \) in this case, we go ahead to the case \( t = 2 \), and so on. This approach is compatible with maximum likelihood decoding. Of course, in practice this approach would be extremely time-consuming. We will address this issue in the next section and obtain at least a theoretical solution.

2.3 Precomputation with Syndromes as Variables

The disadvantage of our decoding method using Gröbner bases is its extremely long computation time. With currently available algorithms for the computation of Gröbner bases it would be impossible ever to use this decoding method in practice. In fact we have been able, with current software, to decode only very small examples. However, if we take a slightly different viewpoint we can achieve a considerable decrease in decoding time, at the cost of a more complex
initial one-time computation. As early as 1960 Peterson [Pe] showed that, for
a BCH code, it is possible to compute rational functions (essentially the same
rational functions found by using Cramer's rule to find the solution to a matrix
equation $Mx^T = s^T$), which give the coefficients of the error locator poly­
nomial in terms of the syndromes. (This technique, as Wolf [Wo] pointed out,
is related to an algorithm proposed by Prony in 1795 for curve-fitting.) Since
the computational complexity of Peterson's approach was too high for more
than about six errors [ML, pp. 142-145], a less complex method of finding the
error locator, designed by Berlekamp [Be] and later modified by Massey [Ma],
became the accepted method for decoding BCH codes. Recently, Cooper [Co]
suggested using the method of Gröbner bases to calculate polynomials that
give the coefficients of the error locator in terms of the syndromes. In this
section we generalize Cooper's method to any linear code.

As in section 2 of this chapter, we view a linear code $C$ as the affine varie­
ty code $C^\perp(I, L)$, where $L = \langle \bar{f}_1, \ldots, \bar{f}_m \rangle$ and $V(I_q) = \{P_1, \ldots, P_n \} \subseteq \mathbb{F}_q^*$. We
assume that the set $\{g_1, \ldots, g_t\}$ of polynomials of $\mathbb{F}_q[x_1, \ldots, x_s]$ that generate
$I_q$ specifically includes the polynomials $x_i^q - x_i, \ldots, x_s^q - x_s$ with which we
construct $I_q$ from $I$.

We view not only the $e_i$ and $P_{ki}, i = 1, \ldots, t$, as variables, but also the
syndrome $(s_1, \ldots, s_m)$. In our earlier work the syndrome was a fixed vector in
$\mathbb{F}_q^m$. Now we replace the values $S_1, \ldots, S_m$ with indeterminates and define the
ideal $\mathcal{E} \subseteq \mathbb{F}_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{ts}, e_t]$ as follows.

**Definition 2.3.1**

\[
\mathcal{E} = \mathcal{E}(t) := \left( \sum_{j=1}^t e_j f_i(x_{j1}, \ldots, x_{js}) - s_i, e_i^{q-1} - 1, g_h(x_{j1}, \ldots, x_{js}), s_i^{q-1} - s_i \right),
\]

where $h = 1, \ldots, l; i = 1, \ldots, m; j = 1, \ldots, t.$
Note that if we replaced the indeterminates $s_1, \ldots, s_m$ by the coordinates $S_1, \ldots, S_m$ respectively of an actual syndrome $(S_1, \ldots, S_m)$ of an error vector of weight $t$, we would have the generating set for the ideal $\mathcal{E}(s_1, \ldots, s_m)$ of $F_q[x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{t1}, \ldots, x_{ts}, e_t]$. (The polynomials $S_i^q - S_i$ in that case would be identically zero, since $S_i \in F_q$ for $1 \leq i \leq m$.)

Next, we define two projection maps $\tau$ and $\rho$ on the variety $V(\mathcal{E})$ of the ideal $\mathcal{E}$. We denote an element of $V(\mathcal{E})$ by

$$(S_1, \ldots, S_m, B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t).$$

Define $\tau : V(\mathcal{E}) \longrightarrow F_q^m$ by

$$\tau(S_1, \ldots, S_m, B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) = (S_1, \ldots, S_m).$$

The fiber above $(S_1, \ldots, S_m)$, denoted $\tau^{-1}((S_1, \ldots, S_m))$, is the set of all elements of $V(\mathcal{E})$ having the field elements $S_1, \ldots, S_m$, in that order, as their first $m$ coordinates.

Define $\rho : V(\mathcal{E}) \longrightarrow F_q^{(s+1)t}$ by

$$\rho(S_1, \ldots, S_m, B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) = (B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t).$$

**Proposition 2.3.2** We have $\rho(\tau^{-1}((S_1, \ldots, S_m))) = V(\mathcal{E}(s_1, \ldots, s_m))$, using the notation above.

**Proof.** Suppose $(B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t) \in V(\mathcal{E}(s_1, \ldots, s_m))$. Then, by definition of $V(\mathcal{E}(s_1, \ldots, s_m))$, every polynomial in $\mathcal{E}(s_1, \ldots, s_m)$ vanishes at $(B_{11}, \ldots, B_{1s}, E_1, \ldots, B_{t1}, \ldots, B_{ts}, E_t)$. In particular, this is true for every polynomial in the generating set defining $\mathcal{E}(s_1, \ldots, s_m)$. Therefore, the polynomials

$$\sum_{j=1}^{s} e_j f_i(x_{j1}, \ldots, x_{js}) - s_i e_j^{s-1} - 1, g_h(x_{j1}, \ldots, x_{js}),$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
all vanish when we replace $e_j$ by $E_j, x_{j_1}, \ldots, x_{j_s}$ by $B_{j_1}, \ldots, B_{j_s}$ respectively, and $s_i$ by $S_i$, where $h = 1, \ldots, l; \ i = 1, \ldots, m; \text{and} \ j = 1, \ldots, t$. Moreover, since $S_i \in \mathbb{F}_q$ for $i = 1, \ldots, m$, it is also true that $s_i^t - s_i$ vanishes when we replace $s_i$ by $S_i$ for $i = 1, \ldots, m$. Hence

$$(S_1, \ldots, S_m, B_{i_1}, \ldots, B_{i_t}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t) \in V(\mathcal{E})$$

and is clearly an element of $\tau^{-1}((S_1, \ldots, S_m))$. Since

$$(S_1, \ldots, S_m, B_{i_1}, \ldots, B_{i_t}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t) \in \tau^{-1}((S_1, \ldots, S_m)),$$

$$(B_{1_1}, \ldots, B_{1_s}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t) \in \rho(\tau^{-1}((S_1, \ldots, S_m))).$$

We have shown that

$$V(\mathcal{E}(S_1, \ldots, S_m)) \subseteq \rho(\tau^{-1}((S_1, \ldots, S_m))).$$

To show the reverse inclusion, suppose

$$(B_{1_1}, \ldots, B_{1_s}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t) \in \rho(\tau^{-1}((S_1, \ldots, S_m))).$$

Then $(S_1, \ldots, S_m, B_{i_1}, \ldots, B_{i_t}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t)$ is an element of $V(\mathcal{E})$, which means the polynomials

$$\sum_{j=1}^{t} e_j f_i(x_{j_1}, \ldots, x_{j_s}) - s_i, e_j^{t-1} - 1, g_h(x_{j_1}, \ldots, x_{j_s}),$$

vanish when $S_i$ is substituted for the indeterminate $s_i$, $B_{j_1}, \ldots, B_{j_s}$ is substituted for the indeterminate $x_{j_1}, \ldots, x_{j_s}$ respectively, and $E_j$ is substituted for the indeterminate $e_j$, where $h = 1, \ldots, l; \ i = 1, \ldots, m; \text{and} \ j = 1, \ldots, t$. But that means that $(B_{1_1}, \ldots, B_{1_s}, E_1, \ldots, B_{t_1}, \ldots, B_{t_s}, E_t)$ is an element of $V(\mathcal{E}(S_1, \ldots, S_m))$. Thus

$$\rho(\tau^{-1}((S_1, \ldots, S_m)) \subseteq V(\mathcal{E}(S_1, \ldots, S_m)),$$

and we are done. □
Corollary 2.3.3 To find the error vector of weight \( t \) whose syndrome is \((S_1, \ldots, S_m)\), it suffices to find the first \( s + 1 \) coordinates of all points in \( \rho(\tau^{-1}((S_1, \ldots, S_m))) \).

Proof. Since \( \rho(\tau^{-1}((S_1, \ldots, S_m))) = V(E_{(S_1, \ldots, S_m)}) \) by the theorem we have just proved, we can find the error vector by finding the first \( s + 1 \) coordinates, by Proposition 2.2.13 and Remark 2.2.14. \( \square \)

As before, we plan to find these coordinates using elimination theory with a Gröbner basis for \( E \). Although we have already proved a similar proposition in Section 2.2, we restate the result for this case:

Corollary 2.3.4 For any intersection ideal \( J = E \cap \mathbb{F}_q[X] \), where \( X \) is any subset of the set \( \{s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{i1}, \ldots, x_{is}, e_i\} \).

1. \( J \) is a radical ideal.

2. If \( \pi \) is the projection map on the coordinates of \( V(E) \) corresponding to the variables in the subset \( X \), then \( \pi(V(E)) = V(J) \).

Proof. Since the polynomials \( s_i^t - s_i \) are elements of \( E \), the same arguments that we used in the proof of Lemma 2.2.16 work in this case. \( \square \)

Notation 2.3.5 Let \( <_x \) be any term order on the variables \( x_1, \ldots, x_m \); let \( <_y \) be any term order on the variables \( y_1, \ldots, y_n \); and let \( <_z \) be any term order on the variables \( z_1, \ldots, z_r \).

Define an elimination order \( <_{xy} \) on the variables \( x_1, \ldots, x_m, y_1, \ldots, y_n \), as in Definition 2.2.20, using \( <_x \) on the \( x \) variables \( x_1, \ldots, x_m \) and \( <_y \) on the \( y \) variables \( y_1, \ldots, y_n \), with the \( y \) variables larger than the \( x \) variables.

Similarly, define an elimination order \( <_{yz} \) on the variables \( y_1, \ldots, y_n, z_1, \ldots, z_r \), as in Definition 2.2.20, using \( <_y \) on the \( y \) variables \( y_1, \ldots, y_n \) and \( <_z \) on the \( z \) variables \( z_1, \ldots, z_r \), with the \( z \) variables larger than the \( y \) variables.
Finally, define an elimination order $<$ on the variables $x_1, \ldots, x_m, y_1, \ldots, y_n, z_1, \ldots, z_r$, as in Definition 2.2.20, using $<_x$ on the $x$ variables $x_1, \ldots, x_m$ and $<_y$ on the $y$ and $z$ variables $y_1, \ldots, y_n, z_1, \ldots, z_r$, with the $y$ and $z$ variables larger than the $x$ variables.

**Proposition 2.3.6** With the notation above, $<$ is the same as an elimination order $<'$ on the variables $x_1, \ldots, x_m, y_1, \ldots, y_n, z_1, \ldots, z_r$, as in Definition 2.2.20, using $<_x$ on the $x$ and $y$ variables and $<_z$ on the $z$ variables, with the $z$ variables larger than the $x$ and $y$ variables. Moreover, when restricted to the variables $x_1, \ldots, x_m, y_1, \ldots, y_n$, $<$ is the same as an elimination order using $<_x$ on the $x$ variables and $<_y$ on the $y$ variables with the $y$ variables larger than the $x$ variables.

**Proof.** Let $X_1Y_1Z_1$ and $X_2Y_2Z_2$ be monomials in the variables

$$x_1, \ldots, x_m, y_1, \ldots, y_n, z_1, \ldots, z_r.$$  

We will show that $X_1Y_1Z_1 < X_2Y_2Z_2$ if and only if $X_1Y_1Z_1 <' X_2Y_2Z_2$.

Suppose that $X_1Y_1Z_1 < X_2Y_2Z_2$. By definition of $<$, that means that

1. $Y_1Z_1 <_y Y_2Z_2$, or
2. $Y_1Z_1 = Y_2Z_2$ and $X_1 <_x X_2$.

First consider case (1). If $Y_1Z_1 <_y Y_2Z_2$, then, by definition of $<_y$, either $Z_1 <_z Z_2$, or $Z_1 = Z_2$ and $Y_1 <_y Y_2$.

If $Z_1 <_z Z_2$, then $X_1Y_1Z_1 <' X_2Y_2Z_2$, by definition. If $Z_1 = Z_2$ and $Y_1 <_y Y_2$, then we have $Z_1 = Z_2$ and $X_1Y_1 <_z X_2Y_2$, so $X_1Y_1Z_1 <' X_2Y_2Z_2$.

Next, consider case (2). If $Y_1Z_1 = Y_2Z_2$, then $Y_1 = Y_2$ and $Z_1 = Z_2$. If $X_1 <_x X_2$, then $X_1Y_1 <_z X_2Y_2$, by definition of $<_z$; by definition of $<'$, $Z_1 = Z_2$ and $X_1Y_1 <_x X_2Y_2$ implies $X_1Y_1Z_1 <' X_2Y_2Z_2$.  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Thus, $X_1Y_1Z_1 < X_2Y_2Z_2$ implies $X_1Y_1Z_1 < X_2Y_2Z_2$.

A similar argument shows that $X_1Y_1Z_1 < X_2Y_2Z_2$ implies $X_1Y_1Z_1 < X_2Y_2Z_2$.

To prove the second claim, note that if we restrict $<$ to the $x$ and $y$ variables, we may simply assume that we are dealing with monomials $X_1Y_1Z_1$ and $X_2Y_2Z_2$, where $Z_1 = Z_2 = z_1^0 \cdots z_n^0$. Then, by definition, $X_1Y_1Z_1 < X_2Y_2Z_2$ when $X_1Y_1 <_{xy} X_2Y_2$. Since $<_{xy}$ is an elimination order using $<_x$ on the $x$ variables and $<_y$ on the $y$ variables, with the $y$ variables larger than the $x$ variables, we are done. □

**Remark 2.3.7** We may apply Proposition 2.3.6 inductively to construct term orders that successively eliminate various subsets of a collection of indeterminates. That method gives us the following corollary, which is of interest for our decoding problem.

**Corollary 2.3.8** Let $<_s$ be any term order on the variables $s_1, \ldots, s_m$; let $<_x$ be the elimination order on the variables $x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{ts}, e_t$ that we used in Section 2.2. [Recall that $<_x$ is an elimination order defined using the lexicographic term order $<_1$ on the variables $x_{11}, \ldots, x_{1s}, e_1$, with $x_{11} <_1 x_{12} <_1 \cdots <_1 x_{1s} <_1 e_1$, and any term order $<_2$ on the remaining variables $x_{21}, \ldots, x_{2s}, e_2, \ldots, e_t$. Also recall that $<_x$ is defined so that the variables $x_{21}, \ldots, x_{2s}, e_2, \ldots, x_{ts}, e_t$ are larger than the variables $x_{11}, \ldots, x_{1s}, e_1$.] Now define an elimination order $<_<$ on the variables $s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{ts}, e_t$, using Definition 2.2.20, using $<_s$ and $<_x$ on their respective sets of variables with the variables $x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{ts}, e_t$ larger than the variables $s_1, \ldots, s_m$. Then

1. $<_<$ is an elimination order with the variables $x_{21}, \ldots, x_{2s}, e_2, \ldots, x_{ts}, e_t$ larger than the variables $s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1$.
2. when restricted to \( s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, < \) is an elimination order with the variable \( e_1 \) larger than the variables \( s_1, \ldots, s_m, x_{11}, \ldots, x_{1s} \);

3. when restricted to \( s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, < \) is an elimination order with the variables \( x_{1j}, \ldots, x_{1s}, e_1 \) larger than the variables \( s_1, \ldots, s_m, x_{11}, \ldots, x_{1(j-1)} \), for any \( j, 1 < j \leq s \);

4. when restricted to \( s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, < \) is an elimination order with the variables \( x_{11}, \ldots, x_{1s}, e_1 \) larger than the variables \( s_1, \ldots, s_m \).

**Proof.** Apply Proposition 2.3.6 inductively, recalling (Remark 2.2.21) that the lexicographic term order \(<_1 \) on the variables \( x_{11}, \ldots, x_{1s}, e_1 \), with \( x_{11} <_1 x_{12} <_1 \cdots <_1 x_{1s} <_1 e_1 \) is an elimination order with \( x_{1j}, \ldots, x_{1s}, e_1 \) larger than \( x_{11}, \ldots, x_{1(j-1)} \) for any \( j, 1 < j \leq s \). Claim (4) simply restates, for convenience, the definition of \(<\). \(\square\)

**Remark 2.3.9** As a shortcut, we might say that \(<\) is a sort of lexicographic term order with

\[
\{s_1, \ldots, s_m\} < x_{11} < \cdots < x_{1s} < e_1 < \{x_{21}, \ldots, x_{2s}, e_2, \ldots, x_{11}, \ldots, x_{1s}, e_1\}.
\]

We may now apply Theorem 2.2.24, which we restate here for reference.

**Theorem 2.2.24 [AL, p. 69]** Let \( I \) be a nonzero ideal of

\[
\mathbb{F}_q[x_1, \ldots, x_m, y_1, \ldots, y_n]
\]

and let \(<\) be an elimination order with the \( y \) variables larger than the \( x \) variables. Let \( G \) be a Gröbner basis for \( I \), computed with respect to \(<\). Then \( G \cap \mathbb{F}_q[x_1, \ldots, x_m] \) is a Gröbner basis for the ideal \( I \cap \mathbb{F}_q[x_1, \ldots, x_m] \).

**Corollary 2.3.10** Let \( G \) be a Gröbner basis for the ideal \( \mathcal{E} \) of

\[
\mathbb{F}_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1, \ldots, x_{t1}, \ldots, x_{ts}, e_t],
\]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
computed with respect to the term order $<$, where $\mathcal{E}$ is defined as in Definition 2.3.1 and $<$ is defined as in Corollary 2.3.8. Then

1. $G \cap F_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1]$ is a Gröbner basis for the ideal $\mathcal{E} \cap F_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1s}, e_1]$;

2. $G \cap F_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1(j-1)}]$ is a Gröbner basis for the ideal $\mathcal{E} \cap F_q[s_1, \ldots, s_m, x_{11}, \ldots, x_{1(j-1)}]$, for $1 \leq j \leq s$;

3. $G \cap F_q[s_1, \ldots, s_m]$ is a Gröbner basis for the ideal $\mathcal{E} \cap F_q[s_1, \ldots, s_m]$.

**Proof.** As we showed in Corollary 2.3.8, the term order $<$ is an elimination order in several ways for various sets of variables. Using this fact along with Corollary 2.3.10 above, we have the desired result. □

**Theorem 2.3.11** Let $<$ be the elimination order defined in Corollary 2.3.8, with $\{s_1, \ldots, s_m\} < x_{11} < \cdots < x_{1s} < e_1 < \{x_{21}, \ldots, x_{2s}, e_2, \ldots, x_{t1}, \ldots, x_{ts}, e_t\}$. Let $G$ be a Gröbner basis for $\mathcal{E}$ with respect to $<$. Let $(S_1, \ldots, S_m)$ be the syndrome of an error vector of weight $t$. Then we can find the error vector by substituting $S_i$ for $s_i$, for $i = 1, \ldots, m$, in the polynomials in $G$ and applying elimination theory to the resulting set of polynomials.

**Proof.** To find the error vector it suffices, by Proposition 2.3.2 and Remark 2.2.14, to find those elements of $\pi(V(\mathcal{E}))$ whose first $m$ coordinates are $S_1, \ldots, S_m$, where $\pi$ is projection on the first $m + s + 1$ coordinates of $V(\mathcal{E})$. Consider the elimination ideal

$$I_{s_1, \ldots, s_m, x_{11}} = \mathcal{E} \cap F_q[s_1, \ldots, s_m, x_{11}].$$

By Corollary 2.3.4, $V(I_{s_1, \ldots, s_m, x_{11}})$ is precisely the set of $s+l(m+1)$-tuples occurring as the first $m+1$ coordinates of $V(\mathcal{E})$. By definition, $V(I_{s_1, \ldots, s_m, x_{11}})$ is the
set of all \((m+1)\)-tuples \((T_1, \ldots, T_m, B_{11}) \in \mathbb{F}_q^{m+1}\) for which all the polynomials in \(I_{s_1, \ldots, s_m, x_{11}}\) vanish when the value \(T_i\) is substituted for the indeterminate \(s_i\), \(1 \leq i \leq m\), and the value \(B_{11}\) is substituted for the indeterminate \(x_{11}\). By Theorem 2.2.24, \(H_1 = G \cap \mathbb{F}_q[s_1, \ldots, s_m, x_{11}]\) is a Gröbner basis for \(I_{s_1, \ldots, s_m, x_{11}}\).

In particular, \(H_1\) is a generating set for \(I_{s_1, \ldots, s_m, x_{11}}\), so to find \(V(I_{s_1, \ldots, s_m, x_{11}})\), it is enough to find \(V(H_1)\). We are interested only in those elements of \(V(H_1)\) whose first \(m\) coordinates are \(S_1, \ldots, S_m\). We already know that these values are valid solutions for the variables \(s_1, \ldots, s_m\) respectively. Substituting \(S_i\) for \(s_i\), \(i = 1, \ldots, m\), in the polynomials of \(H_1\), we obtain a set of univariate polynomials in the variable \(x_{11}\). Solving for \(x_{11}\) by any convenient (or inconvenient) method, we obtain the set of all \((m+1)\)-st coordinates of those elements of \(V(H)\) whose first \(m\) coordinates are \(S_1, \ldots, S_m\). For example, suppose the common roots of the univariate polynomials we obtained are \(R_1, \ldots, R_u\). Then we know that the set of \((m+1)\)-tuples \((S_1, \ldots, S_m, R_1), \ldots, (S_1, \ldots, S_m, R_u)\) is exactly the set of first \(m + 1\) coordinates of those elements of \(V(\mathcal{E})\) whose first \(m\) coordinates are \(S_1, \ldots, S_m\). In other words, this is precisely the set of first \(m + 1\) coordinates of the elements of \(\tau^{-1}(S_1, \ldots, S_m) \subseteq V(\mathcal{E})\), where \(\tau\) is projection on the first \(m\) coordinates, as defined following Definition 2.3.1.

Next, we use these \((m+1)\)-tuples, one by one, as we did in the proof of Theorem 2.2.24, to make substitutions for the variables \(s_1, \ldots, s_m, x_{11}\) in the polynomials of \(H_2 = G \cap \mathbb{F}_q[s_1, \ldots, s_m, x_{11}, x_{12}]\), which, by Theorem 2.2.24, is a Gröbner basis for the ideal \(I_{s_1, \ldots, s_m, x_{11}, x_{12}} = \mathcal{E} \cap \mathbb{F}_q[s_1, \ldots, s_m, x_{11}, x_{12}]\). We obtain, each time, a collection of univariate polynomials in the variable \(x_{12}\), whose common roots we find. Since there are \(u\) \((m+1)\)-tuples, we perform this procedure \(u\) times, each time finding the set of \((m+2)\)-nd coordinates of those elements of \(V(\mathcal{E})\) whose first \((m+1)\) coordinates are \(S_1, \ldots, S_m, R_i\).
for \( i = 1, \ldots, u \). In fact, we are following essentially the same procedure that we used in the proof of Theorem 2.2.25. However, rather than finding all elements of \( \pi(V(\mathcal{E})) \), we restrict the process to those with \( S_1, \ldots, S_m \) as the first \( m \) coordinates.

As before, because of the elimination properties of the term order \( < \), which we constructed for the express purpose of having those properties, we are able to continue the process above until we have found all the coordinates of the elements of \( V(I_{S_1, \ldots, S_m, z_1, \ldots, z_1, e_1}) \) whose first \( m \) coordinates are \( S_1, \ldots, S_m \). Therefore, by our previous results, we have found the error points and their associated error values, and so we have found the error vector. \( \square \)

It is important to note that the results above are true only when the syndrome \( (S_1, \ldots, S_m) \) is actually a syndrome of an error vector of weight exactly \( t \). As in Section 2.2, when we computed the error locator ideal \( \mathcal{E}(S_1, \ldots, S_m)(t) \) for a specific syndrome \( (S_1, \ldots, S_m) \), there will be no common zeros for the polynomials of \( \mathcal{E} \) if \( (S_1, \ldots, S_m) \) is the syndrome of an error vector of weight greater than \( t \). This allows us, given a syndrome of an unknown error vector, to use maximum likelihood decoding. That is, if the syndrome is nonzero, we attempt to decode by assuming that an error of weight 1 has occurred. Only if there are no solutions in that case do we go on to attempt to decode an error of weight 2, and so on. Here is an example.

**Example 2.3.12** Binary five-times repetition code

The five-times repetition code \( C \) over \( \mathbb{F}_2 \) encodes the binary digit 0 to the codeword 00000 and the binary digit 1 to the codeword 11111. Clearly, \( C \) has minimum distance 5 and can therefore correct 2 errors. A parity check matrix
for $C$ is

$$M = \begin{pmatrix}
1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1
\end{pmatrix}$$

which can be expressed as the polynomial evaluation matrix

$$M = \begin{pmatrix}
xy + xz + y + z & (001) & (010) & (011) & (100) & (101) \\
\end{pmatrix}$$

so we may regard $C$ as the affine variety code $C^L(I,L)$, where $I = (xyz + xy + xz + yz + x + y + z + 1, xyz + xy, xyz) \subseteq \mathbb{F}_2[x, y, z]$ and $L$ is the linear subspace of $\mathbb{F}_2[x, y, z]/I_q$ generated by the polynomials $xy + xz + y + z, xy + y, xy + xz + yz + x,$ and $xy + x$. Note that

$$I_q = (xyz + xy + xz + yz + x + y + z + 1, xyz + xy, xyz, x^2 - x, y^2 - y, z^2 - z)$$

and that

$$V(I_q) = \{(001), (010), (011), (100), (101)\}.$$ 

Note also that we have conveniently chosen the points of $V(I_q)$ so that the evaluation of each point as a binary number corresponds to its location in the point order (e.g. the point (100) is in position $4 = 100_2$ in the point order).

Since we want to correct up to two errors we must compute Gröbner bases for the ideals $\mathcal{E}(1)$ and $\mathcal{E}(2)$ corresponding to errors of weights 1 and 2 respectively.
A Gröbner basis for $E(1)$, computed by the software Gb [Fau], is
\[
\begin{align*}
s_4 + x, \\
x^2 + x, \\
s_2 + y \\
xy \\
y^2 + y, \\
s_1 + x + yz + 1, \\
s_3 + y + z + 1, \\
xz + x + yz + y + z + 1, \\
z^2 + z,
\end{align*}
\]

Let us first see what happens when we decode using the syndrome of an error vector of weight 1. For example, suppose the error vector is 00010. Multiplying the transpose of this vector on the left by the parity check matrix $M$, we see that the syndrome of this error vector is (0011); i.e., $S_1 = 0, S_2 = 0, S_3 = 1,$ and $S_4 = 1$. We first try to find the elements of $V(E(1))$ using elimination theory.

By Theorem 2.3.11, the first coordinate of the error point is exactly the set (containing one element) of common solutions to the first group of polynomials of the Gröbner basis above for $E(1)$, when $S_1 = 0, S_2 = 0, S_3 = 1,$ and $S_4 = 1$. Making the substitution $s_4 = 1$, which is the only substitution required in this case, we have
\[
\begin{align*}
s_4 + x & \mapsto 1 + x \\
x^2 + x & \mapsto x^2 + x.
\end{align*}
\]
Since any element of $F_2$ satisfies $x^2 + x$, but only the element 1 of $F_2$ satisfies $x + 1$, we have $x = 1$. We now know that $s_1 = 0, s_2 = 0, s_3 = 1, s_4 = 1,$ and $x = 1$. We substitute the known values for these variables into the polynomials in the second group, which are polynomials in $s_1, s_2, s_3, s_4, x,$ and $y$. We then solve for $y$:
\[
\begin{align*}
s_2 + y & \mapsto 0 + y = y \\
xy & \mapsto y \\
y^2 + y & \mapsto y^2 + y.
\end{align*}
\]
Since any element of $\mathbb{F}_2$ satisfies $y^2 + y$, but only 0 satisfies the polynomial $y$, we have $y = 0$. The information we have so far is $s_1 = 0, s_2 = 0, s_3 = 1, s_4 = 1, x = 1, y = 0$. Making these substitutions in the third group of polynomials, those in the variables $s_1, s_2, s_3, s_4, x, y$, and $z$, we get

\begin{align*}
s_1 + x + yz + 1 & \rightarrow 0 + 1 + 0 + 1 = 0, \\
s_3 + y + z + 1 & \rightarrow 1 + 0 + z + 1 = z, \\
xz + x + yz + y + z + 1 & \rightarrow z + 1 + 0 + z + 1 = 0, \\
z^2 + z & \rightarrow z^2 + z.
\end{align*}

Since any element of $\mathbb{F}_2$ satisfies the polynomials 0 and $z^2 + z$, but only 0 is a root of $z$, we have $z = 0$. We now know all three coordinates of the error point. The error point is (100), for as we have seen, $x = 1, y = 0$, and $z = 0$. We do not need to go on to solve for the error value, because we know it is a nonzero element of $\mathbb{F}_2$, and so it must be 1. Thus the error vector is 00010. Since the variety $V(\mathcal{E}(1))$ is nonempty, we stop. We do not continue on to find elements of $V(\mathcal{E}(2))$, because we already have our decoding solution.

Suppose next that we have received an error vector of weight 2, for example the error vector 11000. The syndrome of this error vector is (0100). That is, $S_1 = 0, S_2 = 1, S_3 = 0, \text{ and } S_4 = 0$. As would be the case in an actual decoding problem, we assume that we do not know the weight of the error vector. Therefore, we first try to find solutions to the polynomials in $\mathcal{E}(1)$. As before, we begin by substituting the known syndrome values for their respective variables in the polynomials in the first group of the Gröbner basis for $\mathcal{E}(1)$. We obtain

\begin{align*}
s_4 + x & \rightarrow x, \\
x^2 + x & \rightarrow x^2 + x.
\end{align*}

We therefore have $x = 0$. Next, we substitute known values for $s_1, s_2, s_3, s_4$, and $x$ into the polynomials of the second group:

\begin{align*}
s_2 + y & \rightarrow 1 + y, \\
x y & \rightarrow 0, \\
y^2 + y & \rightarrow y^2 + y.
\end{align*}
Thus, we have $y = 1$. Continuing, we substitute known values into the third group of polynomials.

$$s_1 + x + yz + 1 \quad \Rightarrow \quad 0 + 0 + z + 1 = z + 1,$$
$$s_3 + y + z + 1 \quad \Rightarrow \quad 0 + 1 + z + 1 = z,$$
$$xz + x + yz + y + z + 1 \quad \Rightarrow \quad 0 + 0 + z + 1 + z + 1 = 0,$$
$$z^2 + z \quad \Rightarrow \quad z^2 + z.$$

Note that the polynomial $z + 1$ implies $z = 1$, while the polynomial $z$ implies $z = 0$. Thus there is no solution to this set of polynomials. That means that the syndrome $(0100)$ is not the syndrome of an error vector of weight 1 (or less). So we must start over again, this time using the Gröbner basis for the ideal $\mathcal{E}(2)$. Again using Gb [Fau], we find that a Gröbner basis for that ideal is:

$$s_4^2 + s_4, $$
$$s_4s_3 + s_4s_1 + s_3s_2s_1 + s_3s_2 + s_3s_1 + s_1, $$
$$s_4s_2s_1 + s_4s_2 + s_4s_1 + s_4, $$
$$s_4^2 + s_3, $$
$$s_2^2 + s_2, $$
$$s_1^2 + s_1, $$

$$x_1^2 + x_1, $$
$$x_1s_4 + x_1s_2 + x_1s_1, $$
$$x_1s_3s_2s_1 + x_1s_3s_2 + x_1s_3s_1 + x_1s_3 + s_3s_2s_1 + s_3s_2 + s_3s_1 + s_3, $$

$$y_1^2 + y_1, $$
$$y_1x_1, $$
$$y_1s_4 + x_1s_2 + x_1s_1, $$
$$y_1s_3 + y_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_3s_2 + s_2s_1, $$
$$y_1s_2s_1 + y_1s_1 + s_4s_2 + s_4 + s_2s_1 + s_1, $$

$$z_1^2 + z_1, $$
$$z_1y_1 + z_1x_1 + z_1 + y_1 + 1, $$
$$z_1x_1s_3 + z_1s_3 + z_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_3s_2 + s_3s_1 + s_2s_1 + s_1, $$
$$z_1x_1s_1 + x_1s_3s_1 + x_1s_1, $$
$$z_1s_4 + x_1s_3 + x_1s_2 + s_4s_1 + s_4s_2 + s_3s_1 + s_3s_2 + s_3s_1 + s_3, $$
$$z_1s_3s_1 + z_1s_1 + x_1s_3s_2 + x_1s_3s_1 + x_1s_3 + x_1s_2s_1 + s_3s_2 + s_3 + s_2s_1 + s_1, $$
$$z_1s_2 + y_1s_2 + y_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_4 + s_2s_1 + s_2 + s_1, $$

$$z_2 + z_1 + s_3 + s_2, $$
$$y_2 + y_1 + s_2, $$
$$x_2 + x_1 + s_4. $$

Note that we have arranged the polynomials in five distinct groups, so that this collection of polynomials has a "triangular" form. That is, the polynomials
in the first group are polynomials in the variables $s_1, \ldots, s_4$ alone, those in the second group are polynomials in $s_1, \ldots, s_4$ and $x_1$, those in the third group are polynomials in $s_1, \ldots, s_4, x_1, y_1$, and those in the fourth group are polynomials in $s_1, \ldots, s_4, x_1, y_1, z_1$. Finally, the fifth group contains the remaining polynomials in this Gröbner basis for $E(2)$. This grouping method makes it easy to find the Gröbner bases, guaranteed by Theorem 2.2.24, for each of the elimination ideals we are interested in. In fact, the Gb software [Fau] we used to calculate this Gröbner basis for $E(2)$ produced its listing in this form, except in reverse order, and without the blank lines we have added between the groups to improve readability.

The polynomials in the first group display relationships among the syndromes $s_1, \ldots, s_4$ of the code $C$. While these relationships may be interesting in themselves, we do not need to concern ourselves with them while we are decoding, for we already have a known syndrome, namely the syndrome of the received word we want to decode. In this example that syndrome is (0100); that is, $s_1 = 0, s_2 = 1, s_3 = 0, s_4 = 0$. We can immediately make these substitutions in the polynomials of the second group, yielding

$$x_1^2 + x_1 \quad \rightarrow \quad x_1^2 + x_1$$
$$x_1 s_4 + x_1 s_2 s_1 + x_1 s_2 + x_1 s_1 \quad \rightarrow \quad x_1$$
$$x_1 s_3 s_2 s_1 + x_1 s_3 s_2 + x_1 s_3 s_1 + x_1 s_3 + s_3 s_2 s_1 + s_3 s_2 + s_3 s_1 + s_3 \rightarrow 0.$$

By Theorem 2.3.11, the set of first coordinates of the error points is precisely the set of simultaneous solutions to these three polynomials. Now every element of $F_2$ satisfies the first polynomial $x_1^2 + x_1$ and also the third polynomial 0. However, only one element of $F_2$ is a root of the second polynomial $x_1$, the field element 0. We therefore know that the first coordinate of the error points is 0. Adding this information to what we already know, we have $s_1 = 0, s_2 = 1, s_3 = 0, s_4 = 0, x_1 = 0$. Making these substitutions in the
polynomials of the third group, we obtain

\[
\begin{align*}
y_1^2 + y_1 & \quad \Rightarrow \quad y_1^2 + y_1 \\
y_1x_1 & \quad \Rightarrow \quad 0 \\
y_1s_4 + x_1s_2 + s_4s_2 & \quad \Rightarrow \quad 0 \\
y_1s_3 + y_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_3s_2 + s_2s_1 & \quad \Rightarrow \quad 0 \\
y_1s_2s_1 + y_1s_1 + s_4s_2 + s_4 + s_2s_1 + s_1 & \quad \Rightarrow \quad 0.
\end{align*}
\]

Clearly, every element of \( \mathbf{F}_2 \) satisfies all these polynomials, and so the set of first and second coordinates of the error points, by Theorem 2.3.11, is exactly the set consisting of 00 and 01. We now have two cases:

1. \( s_1 = 0, s_2 = 1, s_3 = 0, s_4 = 0, x_1 = 0, y_1 = 0 \), and
2. \( s_1 = 0, s_2 = 1, s_3 = 0, s_4 = 0, x_1 = 0, y_1 = 1 \).

In case (1), when we substitute the values 0,1,0,0,0,0 for the variables \( s_1, s_2, s_3, s_4, x_1 \), and \( y_1 \), respectively, in the polynomials of the fourth group, we get

\[
\begin{align*}
z_1^2 + z_1 & \quad \Rightarrow \quad z_1^2 + z_1 \\
z_1y_1 + z_1x_1 + z_1 + y_1 + x_1 + 1 & \quad \Rightarrow \quad z_1 + 1 \\
z_1x_1s_3 + z_1s_3 + z_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_3s_2 + s_3s_1 + s_2s_1 + s_1 & \quad \Rightarrow \quad 0 \\
z_1x_1s_1 + x_1s_3s_1 + x_1s_1 & \quad \Rightarrow \quad 0 \\
z_1s_4 + x_1s_3 + x_1s_2 + s_4s_2 + s_4s_1 + s_3s_2s_1 + s_3s_2 + s_3s_1 + s_3 & \quad \Rightarrow \quad 0 \\
z_1s_3s_1 + z_1s_1 + x_1s_3s_2 + x_1s_3s_1 + x_1s_3 + x_1s_2s_1 + s_3s_2 + s_3 + s_2s_1 + s_1 & \quad \Rightarrow \quad 0 \\
z_1s_2 + y_1s_2 + y_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_4 + s_2s_1 + s_2 + s_1 & \quad \Rightarrow \quad z_1 + 1,
\end{align*}
\]

whose solution set is the element 1 of \( \mathbf{F}_2 \), while in case (2), substituting 0,1,0,0,0,1 for the variables \( s_1, s_2, s_3, s_4, x_1 \), and \( y_1 \), respectively, in the polynomials of the fourth group, we obtain

\[
\begin{align*}
z_1^2 + z_1 & \quad \Rightarrow \quad z_1^2 + z_1 \\
z_1y_1 + z_1x_1 + z_1 + y_1 + x_1 + 1 & \quad \Rightarrow \quad 0 \\
z_1x_1s_3 + z_1s_3 + z_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_3s_2 + s_3s_1 + s_2s_1 + s_1 & \quad \Rightarrow \quad 0 \\
z_1x_1s_1 + x_1s_3s_1 + x_1s_1 & \quad \Rightarrow \quad 0 \\
z_1s_4 + x_1s_3 + x_1s_2 + s_4s_2 + s_4s_1 + s_3s_2s_1 + s_3s_2 + s_3s_1 + s_3 & \quad \Rightarrow \quad 0 \\
z_1s_3s_1 + z_1s_1 + x_1s_3s_2 + x_1s_3s_1 + x_1s_3 + x_1s_2s_1 + s_3s_2 + s_3 + s_2s_1 + s_1 & \quad \Rightarrow \quad 0 \\
z_1s_2 + y_1s_2 + y_1s_1 + x_1s_3s_2 + x_1s_2s_1 + s_4 + s_2s_1 + s_2 + s_1 & \quad \Rightarrow \quad z_1,
\end{align*}
\]

whose solution set is the element 0 of \( \mathbf{F}_2 \). We now know the first, second, and third coordinates of the error points. They are (001) and (010). If we were
working over any field other than $F_2$, we would now go on to find the error values associated with these points by substituting in turn every set of known values for the syndrome and the error points into the group of polynomials in the variables $s_1, \ldots, s_4, x_1, y_1, z_1, e_1$ and solving for $e_1$. In our situation, though, since we know the error values must be nonzero elements of $F_2$, we are done, because there is only one such element, the element 1. Thus, the error vector has a 1 in the first coordinate (001) and a 1 in the second coordinate (010), with zeros elsewhere. Checking this solution against the error vector that we started with, we see that we have decoded correctly.
Chapter 3
Minimum Distance Calculation

3.1 Introduction

In this chapter we show how to find the minimum distance of linear codes using Gröbner basis computations. This is a generalization of results of D. Augot [Au], who described a method for finding minimum weight codewords of cyclic codes using Gröbner bases. We showed in the first chapter that any linear code can be represented as $C^+ (I, L)$ for an ideal $I \subseteq \mathbb{F}_q[x_1, \ldots, x_s]$ (for some $s$) and a linear subspace $L$ of $\mathbb{F}_q[x_1, \ldots, x_s]/I_q$. Now we will see that there is a zero-dimensional polynomial ideal $(S_d)$ for which the variety $V((S_d)) \subseteq \mathbb{F}_q^{d+sd}$ can be mapped onto the set of codewords of weight less than or equal to $d$.

Since $(S_d)$ is zero-dimensional, it is possible (using Gröbner bases) to calculate the number $N$ of points in $V(S_d)$. It is also possible to calculate $N_0$, the number of points of $V(S_d)$ that are mapped to the all-zero codeword. The number $N_0$ is independent of the particular code $C$ and depends only on $q$ and on combinatorial properties of the integer $d$. If $C$ is nontrivial, then the difference $N - N_0$ is zero if and only if the minimum distance of $C$ is strictly greater than $d$.

In Section 2 of this chapter, we show how to calculate $N_0$. In Section 3, we discuss methods of finding a lower bound for the minimum distance of $C$. This enables us to minimize the number of Gröbner basis computations required to find $C$'s minimum distance.
3.2 Calculation of $N_0$

The following three examples illustrate many of the concepts we will work with more abstractly in the remainder of this section. In the examples, we consider the dual code $C$ to the ternary Generalized Reed-Muller code $R(1, 2)$. As an affine variety code, $C$ is the code $C^⊥(I, L)$ where $I$ is the ideal generated by zero in $\mathbb{F}_3[x, y]$, $I_q = (x^3 - x, y^3 - y)$, and $L$ is generated by the polynomials in $x$ and $y$ of total degree less than or equal to 1. The variety $V$ consists of all points of the form $(x, y)$ with both $x$ and $y$ elements of $\mathbb{F}_3$. We order the points of $V$ conventionally. A parity check matrix for this code is given by

\[
\begin{pmatrix}
  f_1 = 1 \\
f_2 = x \\
f_3 = y
\end{pmatrix}
\begin{pmatrix}
  P_{x_1} & P_{x_2} & P_{x_3} & P_{x_4} & P_{y_1} & P_{y_2} & P_{y_3} & P_{y_4} \\
  1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
  0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\
  0 & 1 & 2 & 0 & 1 & 2 & 0 & 1
\end{pmatrix}
\]

Let $S_4$ be the following system of polynomials:

\[
\begin{align*}
a_1 \cdot 1 + a_2 \cdot 1 + a_3 \cdot 1 + a_4 \cdot 1 \\
a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 \\
a_1 y_1 + a_2 y_2 + a_3 y_3 + a_4 y_4 \\
a_i^2 - 1, 1 \leq i \leq 4 \\
x_i^3 - x_i, 1 \leq i \leq 4 \\
y_i^3 - y_i, 1 \leq i \leq 4
\end{align*}
\]

(2)

We seek solutions $a_1, \ldots, a_4, x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4$ to this system of polynomials, and we intend to see how such solutions relate to codewords of weight 4 or less.

**Example 3.2.1 A codeword of weight 4**

It is the first three polynomials of $S_4$ that we are most interested in. The other polynomials simply ensure that $a_1, \ldots, a_4, x_1, y_1, \ldots, x_4, y_4$ all lie in $\mathbb{F}_3$, and that $a_1, \ldots, a_4$ are nonzero.
One solution is the following:

\[
\begin{align*}
    a_1 &= 1, a_2 = 2, a_3 = 2, a_4 = 1, \\
    (x_1, y_1) &= (0, 1) = P_2 \\
    (x_2, y_2) &= (1, 0) = P_4 \\
    (x_3, y_3) &= (1, 2) = P_6 \\
    (x_4, y_4) &= (2, 1) = P_8.
\end{align*}
\]

To confirm that this is a solution, we substitute these elements of \(\mathbb{F}_3\) into the appropriate variables in (2), and, doing arithmetic in \(\mathbb{F}_3\), we have

\[
\begin{align*}
    1 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 + 1 \cdot 1 &= 0 \\
    1 \cdot 0 + 2 \cdot 1 + 2 \cdot 1 + 1 \cdot 2 &= 0 \\
    1 \cdot 1 + 2 \cdot 0 + 2 \cdot 2 + 1 \cdot 1 &= 0
\end{align*}
\]

Since the solution points \(P_2, P_4, P_6,\) and \(P_8\) are distinct, this solution corresponds to a codeword \((01022010)\) of weight 4 with \(a_1 = 1\) in coordinate 2 (corresponding to \(P_2\)), \(a_2 = 2\) in coordinate 4 (corresponding to \(P_4\)), \(a_3 = 2\) in coordinate 6 (corresponding to \(P_6\)), \(a_4 = 1\) in coordinate 8 (corresponding to \(P_8\)), and zeros elsewhere. Note that this really is a codeword, since

\[
\begin{pmatrix}
    1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
    0 & 0 & 0 & 1 & 1 & 2 & 2 & 2 \\
    0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 \\
\end{pmatrix}
\begin{pmatrix}
    0 \\
    1 \\
    2 \\
\end{pmatrix}
= 
\begin{pmatrix}
    0 \\
    2 \\
    0 \\
\end{pmatrix}
\]

Therefore, in this case, there is a codeword of weight 4 corresponding to a solution to \(S_4\). Note that the solution

\[
\begin{align*}
    a_1 &= 2, a_2 = 2, a_3 = 1, a_4 = 1, \\
    (x_1, y_1) &= (1, 0) = P_4 \\
    (x_2, y_2) &= (1, 2) = P_6 \\
    (x_3, y_3) &= (0, 1) = P_2 \\
    (x_4, y_4) &= (2, 1) = P_8.
\end{align*}
\]

also corresponds to the same codeword. This solution is a permutation of the first solution in the sense that we have applied a permutation to the \(a_i\) and the
same permutation to the \((x_i, y_i)\). The codeword we find is the same, because the points of the variety \(V\) are always considered in the same fixed order when we seek codewords.

**Example 3.2.2 A codeword of weight 3**

We continue using the code \(C\) of Example 3.2.1. Another solution to \(S_4\) is

\[
\begin{align*}
  a_1 &= 1, a_2 = 1, a_3 = 2, a_4 = 2, \\
  (x_1, y_1) &= (0, 0) = P_1, \\
  (x_2, y_2) &= (1, 2) = P_6, \\
  (x_3, y_3) &= (2, 1) = P_8, \\
  (x_4, y_4) &= (2, 1) = P_8.
\end{align*}
\]

Note that the points \((x_3, y_3)\) and \((x_4, y_4)\) are both \(P_8\). Confirming that this is a solution, we have

\[
\begin{align*}
  1 \cdot 1 + 1 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 &= 0, \\
  1 \cdot 0 + 1 \cdot 1 + 2 \cdot 2 + 2 \cdot 2 &= 0, \\
  1 \cdot 0 + 1 \cdot 2 + 2 \cdot 1 + 2 \cdot 1 &= 0.
\end{align*}
\]

To see if there is a codeword associated to this solution, we arrange the solution points in order, collect and sum their coefficients, and obtain the following table:

<table>
<thead>
<tr>
<th>Point</th>
<th>Coefficients</th>
<th>Sum of Coefficients (in (F_3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P_1)</td>
<td>(a_1 = 1)</td>
<td>1</td>
</tr>
<tr>
<td>(P_6)</td>
<td>(a_2 = 1)</td>
<td>1</td>
</tr>
<tr>
<td>(P_8)</td>
<td>(a_3 = 2, a_4 = 2)</td>
<td>1</td>
</tr>
</tbody>
</table>

Using the nonzero entries in the third column (Sum of Coefficients), we see that there is a codeword \((100001010)\) with the sum of \(P_1\)'s coefficients in coordinate 1, the sum of \(P_6\)'s coefficients in coordinate 6, the sum of \(P_8\)'s coefficients in coordinate 8, and zeros elsewhere. This is a codeword of weight 3.

This solution to \(S_4\), then, corresponds to a codeword of weight less than 4. Note that there are two coefficients, \(a_3\) and \(a_4\), belonging to the point \(P_8\), and one coefficient each, \(a_1\) and \(a_2\), belonging to the points \(P_1\) and \(P_6\) respectively.
The numbers 2, 1, and 1 of coefficients belonging to the distinct points of the solution constitute a partition of the integer 4. That is, $2+1+1=4$. We will formally define partitions shortly and use them extensively thereafter in this section.

**Example 3.2.3** A solution associated to the zero codeword

Again, we continue working with the code $C$ of Example 3.2.1. A third type of solution occurs when we select $a_1,\ldots,a_4$ so that $a_1 + a_2 + a_3 + a_4 = 0$ and take $(x_1, y_1) = \ldots = (x_4, y_4)$. A specific example is

$$
\begin{align*}
    a_1 &= 1, a_2 = 1, a_3 = 2, a_4 = 2, \\
    (x_1, y_1) &= (2, 2) = P_9 \\
    (x_2, y_2) &= (2, 2) = P_9 \\
    (x_3, y_3) &= (2, 2) = P_9 \\
    (x_4, y_4) &= (2, 2) = P_9.
\end{align*}
$$

The reader can confirm that this is a solution to $S_4$, and indeed, with these $a_i$ we have a solution no matter which particular $P_i \in \{P_1, \ldots, P_9\}$ we choose for the points $(x_i, y_i), 1 \leq i \leq 4$, as long as they are all the same. If we follow the same procedure we used in Example 3.2.2, displaying our results in tabular form, we have:

<table>
<thead>
<tr>
<th>Point</th>
<th>Coefficients</th>
<th>Sum of Coefficients (in $F_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_9$</td>
<td>$a_1 = 1, a_2 = 1, a_3 = 2, a_4 = 2$</td>
<td>0</td>
</tr>
</tbody>
</table>

The codeword associated to this solution has 0 in coordinate 9 and also zeros elsewhere; thus, it is the all-zero codeword. Note that this solution is associated with the partition 4 of the integer 4. That is, all four coefficients belong to the same point.

Of course, combinations of the three types of examples above can also occur. If instead of $S_4$ we used $S_8$ (which means we are interested in solutions corresponding to codewords of weight 8 or less, so we use polynomials
$\sum_{j=1}^{s} a_j f_i(P_j)$ and corresponding polynomials to force the additional variables to meet the qualifications we require), we could have

\[
\begin{align*}
  a_1 &= 1 \quad (x_1, y_1) = P_1 \\
  a_2 &= 1 \quad (x_2, y_2) = P_6 \\
  a_3 &= 2 \quad (x_3, y_3) = P_8 \\
  a_4 &= 2 \quad (x_3, y_3) = P_8
\end{align*}
\]

\[
\begin{align*}
  a_5 &= 1 \quad (x_5, y_5) = P_9 \\
  a_6 &= 1 \quad (x_6, y_6) = P_9 \\
  a_7 &= 2 \quad (x_7, y_7) = P_9 \\
  a_8 &= 2 \quad (x_8, y_8) = P_9.
\end{align*}
\]

Our table then looks like

<table>
<thead>
<tr>
<th>Point</th>
<th>Coefficients</th>
<th>Sum of Coefficients (in $F_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>$a_1 = 1$</td>
<td>1</td>
</tr>
<tr>
<td>$P_6$</td>
<td>$a_2 = 1$</td>
<td>1</td>
</tr>
<tr>
<td>$P_8$</td>
<td>$a_3 = 2, a_4 = 2$</td>
<td>1</td>
</tr>
<tr>
<td>$P_9$</td>
<td>$a_5 = 1, a_6 = 1, a_7 = 2, a_8 = 2$</td>
<td>0</td>
</tr>
</tbody>
</table>

and, constructing a 9-tuple with the sum of the coefficients of $P_i$ in the $i$-th coordinate (and zeros elsewhere), we see that this solution corresponds to the codeword of weight 3 that we have already found.

We next define more rigorously the procedure we have been using to find the codewords associated to a solution to $S_d$.

First, here are a few technical results that we will need later.

**Definition 3.2.4** Fix a finite field $F_q$. Define $z_i$ to be the number of ways to choose $a_1, \ldots, a_i$ from $F_q \setminus \{0\}$ so that $a_1 + \cdots + a_i = 0$. (Note that order matters: $a_1 = 1, a_2 = 4$ is a different solution to $a_1 + a_2 = 0$ in $F_5$ than $a_1 = 4, a_2 = 1$.)

**Remark** Clearly, $z_1 = 0$.

**Lemma 3.2.5** For $i \geq 2$, $z_i = (q-1)^{i-1} - z_{i-1}$.

**Proof.** There are $(q-1)^{i-1}$ ways to choose $a_1, \ldots, a_{i-1}$ from $F_q \setminus \{0\}$. Of these, $z_{i-1}$ have the property that $a_1 + \cdots + a_{i-1} = 0$: thus there is no way to choose a nonzero $a_i$ so that $a_1 + \cdots + a_{i-1} + a_i = 0$. On the other hand, if $a_1 + \cdots + a_{i-1} \neq 0$, there is exactly one way to choose a nonzero $a_i$ (namely...
\[ a_i = -(a_1 + \cdots + a_{i-1}) \] so that \( a_1 + \cdots + a_{i-1} + a_i = 0. \]

**Remark**  The recurrence given in Lemma 3.2.2 provides an easy way to compute the numbers \( z_i \) using a computer program.

It is easy to obtain a closed form for \( z_i \), but in our work it is more convenient to use the recurrence.

**Definition 3.2.6** For \( \beta \in \mathbb{F}_q \setminus \{0\} \), define \( z_{i, \beta} \) to be the number of ways to choose elements \( a_1, \ldots, a_i \) of \( \mathbb{F}_q \setminus \{0\} \) so that \( a_1 + \cdots + a_i = \beta \).

**Lemma 3.2.7** If \( q > 2 \), then \( z_{i, \beta} > 0 \) for any nonzero \( \beta \in \mathbb{F}_q \), and for any \( i > 0 \).

**Proof.** We use induction on \( i \). If \( i = 1 \), take \( a_1 = \beta \).

Now suppose the lemma holds for \( i = k \), and consider the case \( i = k + 1 \).

We seek \( a_1, \ldots, a_{k+1} \in \mathbb{F}_q \setminus \{0\} \) with \( a_1 + \cdots + a_{k+1} = \beta \). Since \( q \neq 2 \), there is a \( \beta' \in \mathbb{F}_q \setminus \{0\} \), with \( \beta' \neq \beta \). By the induction hypothesis, there are nonzero \( a_1, \ldots, a_k \) with \( a_1 + \cdots + a_k = \beta' \). Take \( a_{k+1} = \beta - \beta' \). Since \( \beta \neq \beta' \), \( a_{k+1} \neq 0 \), and \( a_1 + \cdots + a_k + a_{k+1} = \beta' + (\beta - \beta') = \beta \). \( \square \)

**Remark** When \( q = 2 \), there is only one nonzero element of \( \mathbb{F}_q \), namely 1, so

\[ z_{i, 1} = \begin{cases} 1 & \text{if } i \text{ is odd;} \\ 0 & \text{if } i \text{ is even.} \end{cases} \]

Let \( \{P_1, \ldots, P_n\} = V(I_q) \) for some ideal \( I \) of \( \mathbb{F}_q[x_1, \ldots, x_s] \), and use \( R \) as usual to denote \( \mathbb{F}_q[x_1, \ldots, x_s]/I_q \). Let \( \{g_1, \ldots, g_u\} \) be a generating set for \( I \), with \( g_h \in \mathbb{F}_q[x_1, \ldots, x_s] \) for \( h = 1, \ldots, u \). Then, as usual, \( I_q = (g_1, \ldots, g_u, x_i^q - x_i) \) for \( i = 1, \ldots, s \).

Finally, let \( f_1, \ldots, f_m \) be any elements of \( \mathbb{F}_q[x_1, \ldots, x_s] \).

**Definition 3.2.8** Denote by \( S_d \) the union \( S_{d_1} \cup S_{d_2} \cup S_{d_3} \cup S_{d_4} \) of the following sets \( S_{d_1}, S_{d_2}, S_{d_3}, S_{d_4} \) of polynomials in the variables \( a_1, \ldots, a_d, x_{11}, \ldots, x_{1s} \).
\[ ..., x_{d1}, \ldots, x_{ds} : \]

\[ S_{d_1} := \{ \sum_{j=1}^{d} a_j f_i(x_{j1}, \ldots, x_{js}), i = 1, \ldots, m \}, \]
\[ S_{d_2} := \{ a_j^{i-1} - 1, j = 1, \ldots, d \}, \]
\[ S_{d_3} := \{ g_h(x_{j1}, \ldots, x_{js}), h = 1, \ldots, u, j = 1, \ldots, d \}, \]
\[ S_{d_4} := \{ x_{jk}^q - x_{jk}, j = 1, \ldots, d, k = 1, \ldots, s \}. \]

**Lemma 3.2.9** There are a finite number of points in \( V(S_d) \subseteq \mathbb{F}_q^{d+sd} \).

**Proof.** If \((A_1, \ldots, A_d, b_{11}, \ldots, b_{ls}, b_{d1}, \ldots, b_{ds})\) is in \( V((S_d)) \), then \( A_j \) is in \( \mathbb{F}_q \) and \( b_{jk} \) is in \( \mathbb{F}_q \) for \( j = 1, \ldots, d \) and \( k = 1, \ldots, s \). \( \square \)

**Corollary 3.2.10** The ideal generated by \( S_d \) in

\[ \mathbb{F}_q[a_1, \ldots, a_d, x_{11}, \ldots, x_{ls}, \ldots, x_{d1}, \ldots, x_{ds}] \]

is a zero-dimensional ideal.

**Proof.** This follows from the definition of zero-dimensional ideal [Definition 1.1.3]. \( \square \)

**Remark** We will sometimes use the phrase "solution of \( S_d \)" to refer to a point of \( V(S_d) \).

**Lemma 3.2.10** If

\[ s = (A_1, A_2, \ldots, A_d, b_{11}, \ldots, b_{ls}, b_{21}, \ldots, b_{2s}, \ldots, b_{d1}, \ldots, b_{ds}) \]

is a solution of \( S \), then \((b_{j1}, \ldots, b_{js})\) is a point of \( V(I_q) \) for each \( j, 1 \leq j \leq d \).

**Proof.** Each \( s \)-tuple \((b_{j1}, \ldots, b_{js})\) satisfies all the polynomials in \( I_q \), because \( b_{j1}, \ldots, b_{js} \) satisfy the polynomials in \( S_{d_4} \) and \( S_{d_4} \), which together generate \( I_q \). \( \square \)

**Definition 3.2.11** (Notation) For a solution

\[ s = (A_1, A_2, \ldots, A_d, b_{11}, \ldots, b_{ls}, b_{21}, \ldots, b_{2s}, \ldots, b_{d1}, \ldots, a_{ds}) \]
of $S$, where
\[(b_{j_1}, \ldots, b_{j_d}) = P_j \in \{P_1, \ldots, P_n\} = V(I_q),\]
write
\[s = ((A_1, A_2, \ldots, A_d), (P_{l_1}, P_{l_2}, \ldots, P_{l_d})).\]
We will call $A_j$ the coefficient of the point $P_{l_j}$.

**Example 3.2.12.** The solution
\[
\begin{align*}
  a_1 &= 1, a_2 = 2, a_3 = 2, a_4 = 1, \\
  (x_1, y_1) &= (0, 1) = P_2 \\
  (x_2, y_2) &= (1, 0) = P_4 \\
  (x_3, y_3) &= (1, 2) = P_6 \\
  (x_4, y_4) &= (2, 1) = P_8.
\end{align*}
\]
of Example 3.2.1 would be written $((1, 2, 2, 1), (P_2, P_4, P_6, P_8))$. The first point
$P_{l_1}$ in the 4-tuple of points $(P_2, P_4, P_6, P_8)$ is $P_2$ and its coefficient is the first
entry $A_1$ in the 4-tuple $(1, 2, 2, 1)$ of coefficients, namely 1. Thus, the coefficient
of $P_{l_1}$ is 1.

Now let $C = C^I(I, L)$ denote the affine variety code contained in $F_q^n$,
with $L$ the linear subspace of $F_q[x_1, \ldots, x_d]/I_q$ generated by the polynomials
$f_1, \ldots, f_m$ and $P_1, \ldots, P_n$ a fixed ordering of the points of $V = V(I_q)$.

Denote by $T$ the set of solutions of $S_d$. An element of $T$ looks like
\[s = ((A_1, \ldots, A_d), (P_{l_1}, \ldots, P_{l_d})) \in (F_q^n)^d \times V^d.\]
(We are giving $V(S_d)$ a new name $T$ for when we think of it as being contained
in the set $(F_q^n)^d \times V^d$.)

Define a set map $\gamma : T \rightarrow F_q^n$ as follows.

First note that $F_q^n$ is an $F_q$-vector space with canonical basis $\{\xi_i : 1 \leq i \leq n\}$, where $\xi_i$ is the n-tuple having 1 in the i-th coordinate and zeros elsewhere.

Next note that each of the points $P_{l_i}$, $1 \leq i \leq d$, is one of the points $P_1, \ldots, P_n$ of $V(I_q)$. So $l_j$ is an integer between 1 and $n$ inclusive, but the $l_j$
need not all be distinct. The idea of the map $\gamma$ is to collect and sum all the coefficients $A_{ij}$ belonging to each distinct $P_i$.

Therefore, we define $\gamma$ by

$$\gamma((A_1, \ldots, A_d), (P_1, \ldots, P_d)) = \sum_{j=1}^{d} A_j \xi_{ij}.$$  

**Example 3.2.13** Suppose $V(\mathbb{F}_q) = \{P_1, P_2, \ldots, P_5\}$, and suppose

$$((4,4,1,1,2),(P_5,P_3,P_2,P_5,P_3))$$

is a solution to $S_d$, with $\mathbb{F}_q = \mathbb{F}_5$. Then $\gamma$ maps this solution to

$$4(00001) + 4(00100) + 1(01000) + 1(00001) + 2(00100) = (01100).$$

Note especially that the coefficients $A_1 = 4$ and $A_4 = 1$ of $P_3$ sum to 0, so the fifth coordinate of the image under $\gamma$ of this solution is 0.

**Proposition 3.2.14** With the notation and definitions of the preceding paragraphs, and for $q > 2$, the image of $\gamma$ in $\mathbb{F}_q^n$ is the set of codewords of $C$ having weight less than or equal to $d$.

**Proof.** Suppose $s = ((A_1, \ldots, A_d), (P_1, \ldots, P_d))$ is a solution of $S_d$. We have already seen that the $P_{ij}$, $1 \leq j \leq d$, are points of $V(\mathbb{F}_q)$. Since this point of $V(S_d)$ must satisfy the polynomials of $S_{dq}$, $A_j^{q-1} - 1 = 0$ for $j = 1, \ldots, d$, so each $A_j \in \mathbb{F}_q \setminus \{0\} \subseteq \mathbb{F}_q$. Therefore, the sum of any subset of the elements of \{\$A_1, \ldots, A_d\$\} is an element of $\mathbb{F}_q$.

If

$$\gamma((A_1, \ldots, A_d), (P_1, \ldots, P_d)) = \sum_{j=1}^{d} A_j \xi_{ij} = (0, \ldots, 0),$$

we are done, because the all-zero vector is always a codeword. Suppose $\sum_{j=1}^{d} A_j \xi_{ij} \neq (0, \ldots, 0)$. Since there are at most $d$ distinct $\xi_{ij}$ contributing to the sum, $\sum_{j=1}^{d} A_j \xi_{ij}$ is a vector of $\mathbb{F}_q^n$ of weight at most $d$. To see that this
vector is a codeword, note that we may write \( \sum_{j=1}^{d} A_j \xi_j \) as \( \sum_{i=1}^{n} c_i \xi_i \), where 
\[ c_i = \sum_{j=i}^{d} A_j \] if \( l_j = i \) for any \( j, 1 \leq j \leq d \), and \( c_i = 0 \) otherwise. Then, for any polynomial \( f \in \{f_1, \ldots, f_m\} \), \( \sum_{i=1}^{n} c_i f(P_i) \) may be rewritten as 
\[
\sum_{c_i \neq 0} \sum_{l_j=i} A_j f(P_i),
\]
(1)
(leaving out the terms where \( c_i = 0 \)). Distributing \( f(P_i) \) over the second sum, we see, since there are \( d \) coefficients \( A_j \), that equation (1) can be rewritten as 
\[
A_1 f(P_1) + \cdots + A_d f(P_d) = 0,
\]
because the point \( ( (A_1, \ldots, A_d), (P_1, \ldots, P_d) ) \) is a solution of \( S_d \). Since this is true for any \( f \) in \( \{f_1, \ldots, f_m\} \), \( \sum_{j=1}^{d} A_j \xi_j \) is in the orthogonal complement of \( L \) in \( \mathbb{F}_q^n \), so is a codeword of \( C \).

Conversely, suppose \( c = (c_1, \ldots, c_n) \) is a codeword of \( C \) of weight less than or equal to \( d \). If the weight of \( c \) is exactly \( d \), then there are nonzero coordinates \( c_{i_1} = A_1, \ldots, c_{i_d} = A_d \) such that 
\[
A_1 f_i(P_1) + \cdots + A_d f_i(P_d) = 0
\]
for all \( f_i \) with \( 1 \leq i \leq m \). Since the \( A_j \) are nonzero elements of \( \mathbb{F}_q \), they satisfy the polynomials of \( S_{d_1} \) and the points \( P_{i_j}, 1 \leq j \leq d \), satisfy the polynomials of \( S_{d_1} \) and \( S_{d_4} \), since they are points of \( V(I_d) \). Thus
\[
s := ( (A_1, \ldots, A_d), (P_1, \ldots, P_d) ) \in V(S_d),
\]
and \( \gamma \) maps this solution \( s \) of \( S_d \) to the codeword \( c \).

If \( c \) has weight \( w \) less than \( d \), then there are exactly \( w \) nonzero coordinates \( c_{i_1}, \ldots, c_{i_w} \) of \( c \), with \( \sum_{j=1}^{w} c_{i_j} f(P_{i_j}) = 0 \), for all \( f \in \{f_1, \ldots, f_m\} \). Let 
\[
A_1 = c_{i_1}, \ldots, A_{w-1} = c_{i_{w-1}},
\]
and choose \( A_w, A_{w+1}, \ldots, A_d \) from \( \mathbb{F}_q \setminus \{0\} \) such that \( A_w + A_{w+1} + \cdots + A_d = c_i \).

By Lemma 3.2.7, there is at least one way to do this. Then, for any \( f \) in \( \{ f_1, \ldots, f_m \} \),

\[
0 = c_1 f(P_{i_1}) + \cdots + c_{i_w-1} f(P_{i_{w-1}}) + c_{i_w} f(P_{i_w}) \\
= A_1 f(P_{i_1}) + \cdots + A_{i_w-1} f(P_{i_{w-1}}) + c_{i_w} f(P_{i_w}) \\
= A_1 f(P_{i_1}) + \cdots + A_{i_w-1} f(P_{i_{w-1}}) + A_w f(P_{i_w}) + \cdots + A_d f(P_{i_w}),
\]

since \( A_w + \cdots + A_d = c_i \). Thus

\[
((A_1, \ldots, A_{i_w-1}, A_w, \ldots, A_d), (P_{i_1}, \ldots, P_{i_{w-1}}, P_{i_w}, \ldots, P_{i_w}))
\]

is a solution of \( S_d \) (where the point \( P_{i_w} \) occurs in positions \( w \) through \( d \)). Note that \( \gamma \) maps this solution to the codeword \( c \).

\[\square\]

**Remark** The proposition is not true for \( q = 2 \). In that case, unless the minimum distance of the code is 1, the image of \( \gamma \) does not include codewords of weights \( d - (2k + 1) \), for any integer \( k \). So if \( d \) is even, the image of \( \phi \) contains only even-weight codewords, and if \( d \) is odd, the image of \( \phi \) contains only odd-weight codewords. For example, for a binary code of length 5, suppose there is a codeword of weight 4, and no codeword of weight 5. Then all four nonzero coordinates of the weight-4 codeword must be 1, because 1 is the only nonzero element of \( \mathbb{F}_2 \). A solution \( s \) to \( S_5 \) has to have \( A_1 = A_2 = \cdots = A_5 = 1 \). Four of the five \( A_i \) must be coefficients of the points associated with the nonzero coordinates of the weight-4 codeword. The other coefficient, which is also 1, must be associated to some point. It cannot be a point different from the four associated with the codeword, for that would mean the code has a codeword of weight 5. But it cannot be one of the four points associated with the weight-4 codeword, either, because then the sum of the coefficients associated with that point is zero, which implies the code has a codeword of weight 3, three of whose nonzero coordinates match those of a codeword of weight 4 (so the minimum distance of the code is 1). However, we may modify our procedures for finding
minimum distance to take this difficulty into account. We will discuss this later; for now, we assume $q \neq 2$.

**Corollary 3.2.15** If the image of $\gamma$ is not equal to $(0, \ldots, 0)$, then the minimum distance of $C$ is less than or equal to $d$.

Next, we count the number of elements of $T$ that $\gamma$ maps to $(0, \ldots, 0)$. We have already seen that the set $T = V(S_d)$ is finite [Lemma 3.2.9], so it makes sense to talk about counting elements of $\gamma^{-1}(0, \ldots, 0)$.

**Definition 3.2.16** [St, pp. 28-29] For a positive integer $d$, define a partition of $d$ to be a sequence $\lambda = (\lambda_1, \ldots, \lambda_k)$, (where each $\lambda_i$ is a nonnegative integer) such that $\sum \lambda_i = d$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$. It is usual to ignore trailing zeros. For example, we consider $(3, 2, 2, 1)$ and $(3, 2, 2, 1, 0, 0)$ the same partition of 8. The nonzero terms are called the parts of $\lambda$. If there are $k$ nonzero terms, we say that $\lambda$ has $k$ parts. A convenient way to write a partition $\lambda$, where $\lambda$ has $r_i$ parts equal to $i$, is $\lambda = (1^{r_1}, 2^{r_2}, \cdots)$, where we may omit the exponent 1 and terms with $r_1 = 0$. For example,

$$(5, 5, 5, 4, 2, 2, 2, 1) = (1^5, 2^4, 3^0, 4^1, 5^3) = (1, 2^4, 4, 5^3).$$

**Example 3.2.17** The partitions of 5 are

- $(5)$
- $(4, 1)$
- $(3, 2)$
- $(3, 1, 1) = (1^2, 3)$
- $(2, 2, 1) = (1^2, 2^2)$
- $(2, 1, 1, 1) = (1^3, 2)$
- $(1, 1, 1, 1, 1) = (1^5)$

**Definition 3.2.18** Let $\lambda = (d_1, \ldots, d_k)$ be a partition of $d$ with $k$ parts. We say a $d$-tuple $u$ of $V^d$ is associated to $\lambda$ if there are $k$ distinct points $P_{j_1}, \ldots, P_{j_k}$ of $V$ such that $P_{j_w}$ occurs $d_{w}$ times as a coordinate of $u$, $1 \leq w \leq k$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Example 3.2.19  The partition $\lambda = (3,2,2,1)$ of 8 has 4 parts. For $V = \{P_1, \ldots, P_8\}$, the 8-tuple $u = (P_5, P_3, P_5, P_3, P_1, P_3, P_4)$ is associated to $\lambda$ because there are 4 distinct points $P_5, P_3, P_1, P_4$ of $V$ with $P_5$ occurring 3 times as a coordinate of $u$, $P_3$ occurring 2 times, $P_1$ occurring 2 times, and $P_4$ occurring 1 time.

Remark 3.2.20  Note that each element of $V^d$ is associated to exactly one partition of $d$.

Proposition 3.2.21  Let $V = \{P_1, \ldots, P_n\}$. Let $\lambda = (\lambda_1^{r_1}, \ldots, \lambda_d^{r_d})$ be a partition of $d$, and let $k = r_1 + \cdots + r_d$ be the number of parts of $\lambda$. Then there are exactly

$$\binom{n}{k}(k!/r_1! \cdots r_d!)(d!/(\lambda_1!)^{r_1} \cdots (\lambda_d!)^{r_d})$$

elements of $V^d$ associated to $\lambda$.

Proof.  There are $\binom{n}{k}$ ways to choose $k$ distinct points $P_{j_1}, \ldots, P_{j_k}$ from the $n$ points of $V$. There are $k!/r_1! \cdots r_d!$ ways to assign the parts of $\lambda$ to the selected $k$ points. Let $u$ be a point of $V^d$ such that each $P_{j_w}$ occurs its assigned number of times for $w = 1, \ldots, k$. Then all the elements of $V^d$ associated to $\lambda$ are obtained by permuting the coordinates of $u$ in every possible way. There are $d!/(\lambda_1!)^{r_1} \cdots (\lambda_d!)^{r_d}$ distinct ways to do this. Using the multiplication principle, we obtain the result. $\Box$

Lemma 3.2.22  Let $s = ((A_{i_1}, \ldots, A_d), (P_{j_1}, \ldots, P_{j_d})) \in T$. Then $\gamma(s) = (0, \ldots, 0)$ if and only if the sum of the coefficients $A_{i_j}$ associated to the point $P_j$ is zero for $1 \leq j \leq n$.

Proof.  This follows immediately from the definition of the map $\gamma$ and the fact that the basis elements $\xi_i$ of $F_q^n$ are linearly independent. $\Box$
Theorem 3.2.23 For $\gamma$ defined as in Proposition 3.2.14,

$$|\gamma^{-1}(0, \ldots, 0)| = \sum_{\lambda \in \Lambda_d} (z_{\lambda_1})^{r_1} \cdots (z_{\lambda_l})^{r_l}\binom{n}{k}(k!/r_1! \cdots r_l!)(d!/\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l},$$

where $\Lambda_d$ is the set of all partitions of $d$, and where a partition $\lambda = (\lambda_1^{r_1}, \ldots, \lambda_l^{r_l})$, with the number of parts of $\lambda$ equal to $k = r_1 + \cdots + r_l$.

Proof. Let $v = ((a_1, \ldots, a_d), u) \in (F_q^x)^d \times V^d$. Suppose $u$ is associated to the partition $\lambda = (\lambda_1^{r_1}, \ldots, \lambda_l^{r_l})$. Rewrite $\lambda = (d_1, \ldots, d_k)$ with $k = r_1 + \cdots + r_l$.

Then there are $k$ distinct points $P_{j_1}, \ldots, P_{j_k}$ of $V$ occurring as coordinates of $u$, with $P_{j_i}$ occurring $d_i$ times. Let $A_{i_1}, \ldots, A_{i_d}$ be the coefficients of $P_{j_i}$ in $v$. Note that all these coordinates are nonzero. By Lemma 3.2.20, $\gamma(v) = 0$ if and only if $\sum_{w=1}^{d_i} A_{i,w} = 0$ for $i = 1, \ldots, k$.

By Definition 3.2.4, there are exactly $z_{d_i}$ ways to choose nonzero $a_{i,w}$, $w = 1, \ldots, d_i$, from $F_q$ so that $a_{i_1} + \cdots + a_{i_d} = 0$. Hence there are exactly $z_{d_1}z_{d_2} \cdots z_{d_k} = (z_{\lambda_1})^{r_1} \cdots (z_{\lambda_l})^{r_l}$ ways to choose the $a_{i,w}$, $1 \leq w \leq d_i$, $1 \leq i \leq k$, so that $\gamma((a_1, \ldots, a_d), u) = 0$.

By Proposition 3.2.21, there are $\binom{n}{k}(k!/r_1! \cdots r_l!)(d!/\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l}$ elements $u \in V^d$ associated to $\lambda$. Thus, there are

$$(z_{\lambda_1})^{r_1} \cdots (z_{\lambda_l})^{r_l}\binom{n}{k}(k!/r_1! \cdots r_l!)(d!/\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l}$$

elements $((a_1, \ldots, a_d), u)$ of $\gamma^{-1}(0)$ with $u$ associated to $\lambda$.

Since each element $u$ of $V^d$ is associated to exactly one partition $\lambda$ of $d$, we may take the sum over all partitions $\lambda$ of $d$ to obtain

$$|\gamma^{-1}((0, \ldots, 0))| = \sum_{\lambda \in \Lambda_d} (z_{\lambda_1})^{r_1} \cdots (z_{\lambda_l})^{r_l}\binom{n}{k}(k!/r_1! \cdots r_l!)(d!/\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l},$$

as we wanted. $\square$

Definition 3.2.24 Define $N_0 = N_0(d)$ to be $|\gamma^{-1}((0, \ldots, 0))|$ for the system $S_d$.  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Example 3.2.25 Suppose $C$ is a code of length $n = 7$ over $F_8$. (So $q = 8$). An example of such a code is the Reed-Solomon code whose parity check matrix is given by the polynomial evaluation matrix

$$
\begin{pmatrix}
\alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 = 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 \\
x & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \\
x^2 & \alpha^2 & (\alpha^2)^2 & (\alpha^3)^2 & (\alpha^4)^2 & (\alpha^5)^2 & (\alpha^6)^2 & 1
\end{pmatrix}
$$

However, the details of $C$'s parity check matrix are unnecessary to calculate $N_0(d)$, which depends only on $n$, $q$, and $d$. Suppose first that $d = 3$.

The partitions of 3 are

$$
\begin{align*}
(3) \\
(2,1) \\
(1,1,1) = \langle 1^3 \rangle.
\end{align*}
$$

Since $z_1 = 0$, partitions containing a 1 contribute nothing to the sum $N_0$. Thus in this case we need only calculate the term corresponding to the partition 3 of 3.

For $\lambda = (3)$, we have

$$
(z_{\lambda_1})^{r_1} \cdots (z_{\lambda_\nu})^{r_\nu} \binom{n}{k} \frac{k!}{r_1! \cdots r_\nu!} (d! / (\lambda_1!)^{r_1} \cdots (\lambda_\nu!)^{r_\nu}) = (z_3) \frac{\binom{3}{1}}{1} (3!/3!).
$$

Since

$$
z_3 = (q - 1)^2 - z_2 = (q - 1)^2 - (q - 1)^1 + 0 = 7^2 - 7 = 42,
$$

we obtain

$$
N_0(3) = 42 \cdot 7 = 294.
$$

Next suppose that $d = 4$. The partitions of 4 are

$$
\begin{align*}
(4) \\
(3,1) \\
(2,2) = \langle 2^2 \rangle \\
(2,1,1) = \langle 1^2,2 \rangle \\
(1,1,1,1) = \langle 1^4 \rangle.
\end{align*}
$$
As before, those partitions containing a 1 contribute nothing to $N_0(4)$, so for $\lambda = (4)$, we have

$$(z_{\lambda_1})^{r_1} \cdots (z_{\lambda_n})^{r_n} \binom{n}{k} (k! / r_1! \cdots r_l!)(d! / (\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l}) = z_4 \cdot 7,$$

and for $\lambda = (2,2)$, we have

$$(z_{\lambda_1})^{r_1} \cdots (z_{\lambda_n})^{r_n} \binom{n}{k} (k! / r_1! \cdots r_l!)(d! / (\lambda_1!)^{r_1} \cdots (\lambda_l!)^{r_l}) = (z_2)^2 \binom{7}{2} (2! / 2!) (4! / 2!)^2).$$

Since $z_4 = 7^3 - z_3 = 343 - 42 = 301$, and since $z_2 = 7$, we obtain 2107 for the first term and $49 \cdot 21 \cdot 6 = 6174$ for the second.

Their sum, $N_0(4)$, is $2387 + 6174 = 8281$.

We calculate the total number of solutions $N$ to the system $S_d$ using the software package Gb [Fau] that computes a Gröbner basis for $S_d$ and also computes the number of points in $V(S_d))$. We then compare $N$ and $N_0$ to see if the code $C$ has codewords of weight $d$ or less.

The results of the computation are that the number of solutions to the system $S_3$ is $N(3) = 294$ and the number of solutions to the system $S_4$ is $N(4) = 14161$. Since $N_0(3) = 294$, there are no nonzero codewords for this code of weight less than or equal to 3. Since $N(4) = 14161$ and $N_0(4) = 8281$, the difference $N(4) - N_0(4) = 14161 - 8281 = 5880 > 0$ tells us that there are codewords of weight less than or equal to 4. Since we already know there are no codewords of weight less than or equal to 3, we know that all the solutions not counted in $N_0$ belong to codewords of weight 4. Thus their number is the number of all the permutations of the nonzero coefficients of the codewords of weight 4. Hence, there are $(N - N_0)/4! = 5880/24 = 245$ codewords of weight 4. This figure is corroborated by [TV, p. 41].

**Remark** We can, if necessary, modify $S_d$ so that we are looking for solutions associated with a subfield subcode of $C$. This is easy to do by changing the
polynomials $a_j^{q-1} - 1$ to $a_j^{q_0-1} - 1$, where $F_{q_0}$ is a subfield of $F_q$. We would also have to change the way that we calculate $N_0$, by computing the numbers $z_i$ with respect to $q_0$ rather than $q$.

We now return to the case $q = 2$. As we discussed earlier, when $q = 2$, $\gamma(V(S_d))$ is not necessarily the set of all codewords of weight less than or equal to $d$. However, it is certainly true that $\gamma(V(S_d))$ does contain all codewords of weight $d$. Therefore, if we already know that a code $C$ has no codewords of weight strictly less than $d$, then $N(d) > N_0(d)$ means that $C$ has at least one codeword of weight exactly $d$. Our usual method to determine the minimum distance of a code is going to be to find a lower bound $b$ for the minimum distance, then carry out computations for $S_b, S_{b+1}, \ldots$ until we have $N > N_0$.

We can use this same method for $q = 2$.

3.3 A Lower Bound on Minimum Distance

In certain cases we can find a lower bound on the minimum distance of a linear code when it is presented as an affine variety code. Such a bound allows us to reduce the number of ideals $(S_d)$ for which we must compute a Gröbner basis in order to find the code's minimum distance using the methods of Section 3.2.

The material in this section is adapted from parts of [HLP], where it is given in detail for evaluation codes and their duals.

Let $N$ be the set of positive integers $\{1, 2, 3, \ldots\}$, and let $N_0$ be the set of nonnegative integers $\{0, 1, 2, 3, \ldots\}$. Let $R$ be a commutative ring with 1 that contains a field $F$ as a subring. Note that we may regard $R$ as an $F$-vector space; hence $R$ has a basis $\{f_1, f_2, \ldots\}$. (In our case, $R$ is usually $F_q[x_1, \ldots, x_s]$, $F_q[x_1, \ldots, x_s]/I$, or $F_q[x_1, \ldots, x_s]/I_q$.)

Consider a map

$$\rho : R \rightarrow N_0 \cup \{-\infty\}.$$
**Definition 3.3.1** [HLP, p. 20] The map \( \rho \) is called an order function on \( R \) if, for all \( f, g, h \in R \), the following conditions hold:

1. **(O.0)** \( \rho(f) = -\infty \) if and only if \( f = 0 \).
2. **(O.1)** \( \rho(\lambda f) = \rho(f) \) for all nonzero \( \lambda \in F \).
3. **(O.2)** \( \rho(f + g) \leq \max\{\rho(f), \rho(g)\} \), and equality holds when \( \rho(f) < \rho(g) \).
4. **(O.3)** If \( \rho(f) < \rho(g) \) and \( h \neq 0 \), then \( \rho(fh) < \rho(gh) \).
5. **(O.4)** If \( \rho(f) = \rho(g) \), then there exists a nonzero \( \lambda \in F \) such that \( \rho(f - \lambda g) < \rho(g) \).

**Definition 3.3.2** [HLP, p. 20] The map \( \rho \) is called a weight function if it is an order function on \( R \) and in addition satisfies:

6. **(O.5)** \( \rho(fg) = \rho(f) + \rho(g) \).

**Proposition 3.3.3** Suppose \( R \) has an order function \( \rho \). Then there exists a basis \( \{f_i : i \in \mathbb{N}\} \) of \( R \) over \( F \) such that \( \rho(f_i) < \rho(f_{i+1}) \) for all \( i \in \mathbb{N} \).

**Proof.** See [HLP, p. 24, Proposition 3.12]. \( \square \)

**Definition 3.3.4** [HLP, p. 55] Let \( R \) have an order function \( \rho \). Fix a basis \( \{f_i : i \in \mathbb{N}\} \) of \( R \) such that \( \rho(f_i) < \rho(f_{i+1}) \). The number \( l(i, j) \) is the unique \( l \) such that \( \rho(f_if_j) = \rho(f_i) \). The basis elements will be called monomials, and the set of monomials will be denoted by \( M \). If \( f = \sum_{i=1}^{l} \lambda_if_i \) and \( \lambda_j \neq 0 \), then \( \{f_i : \lambda_i \neq 0\} \) is called the support of \( f \). The leading monomial of \( f \) is \( f_j \) and is denoted by \( \text{lm}(f) \).

We gave a review of Gröbner bases in Section 2.2. The following theorem characterizes Gröbner bases, using the notation above.
**Theorem 3.3.5** [HLP, p. 56, Theorem 6.4] Let \( B \) be a finite set in \( R \). Then \( B \) is a Gröbner basis if and only if

\[
\{\text{lm}(f) : f \in (B), f \neq 0\} = \{\text{lm}(bm) : b \in B, b \neq 0, m \in M\}.
\]

**Proof.** See [HLP].

**Definition 3.3.6** [HLP, p. 56] The footprint or \( \Delta \) set of a Gröbner basis \( B \) is defined by

\[
\Delta(B) = M \setminus \{\text{lm}(bm) : b \in B, b \neq 0, m \in M\}.
\]

**Remark** If \( B \) is a Gröbner basis for the ideal \( I \) in \( R \), then the footprint \( \Delta(B) \) is a set of representatives for a basis of \( R/I \).

**Definition 3.3.7** [HLP, p. 23, CLO, p. 73] Let \( w = (w_1, \ldots, w_s) \) be an s-tuple of positive integers. For \( \alpha = (\alpha_1, \ldots, \alpha_s) \in \mathbb{N}_0^s \), let \( X^\alpha = x_1^{\alpha_1} \cdots x_s^{\alpha_s} \) be a monomial of \( F[x_1, \ldots, x_s] \) (where in this case we use the term monomial in the usual sense). The weighted degree \( w\text{deg}(X^\alpha) \) of the monomial \( X^\alpha \) is defined by \( w\text{deg}(X^\alpha) = \sum_{j=1}^s \alpha_j w_j \), and the weighted degree of a nonzero polynomial \( f = \sum \lambda_\alpha X^\alpha \) is defined by \( w\text{deg}(f) = \max\{w\text{deg}(X^\alpha) : \lambda_\alpha \neq 0\} \). The total weighted degree lexicographic order \( <_W \) is defined by \( X^\alpha <_W X^\beta \) if and only if \( w\text{deg}(X^\alpha) < w\text{deg}(X^\beta) \) or \( w\text{deg}(X^\alpha) = w\text{deg}(X^\beta) \) and \( X^\alpha <_L X^\beta \), where \( <_L \) is the lexicographic order of Definition 2.2.19.

**Proposition 3.3.8** Let \( I \) be an ideal in \( F[x_1, \ldots, x_s] \) with Gröbner basis \( B \) with respect to \( <_W \). Suppose that the elements of the footprint of \( I \) have mutually distinct weighted degrees and that every element of \( B \) has two monomials of highest weighted degree in its support. Then there exists a weight function \( \rho \) on \( R = F[x_1, \ldots, x_s]/I \) with the property that \( \rho(\bar{f}) = w\text{deg}(f) \), where \( \bar{f} \) is the coset of \( f \) modulo \( I \), for all polynomials \( f \).
Proof. [HLP, p. 57, Proposition 6.8.]

Thus, for a finite affine variety $V = \{P_1, \ldots, P_n\} \subseteq \mathbb{F}_q^*$, where $V = V(I_q)$ with $I_q$ satisfying the conditions of the preceding proposition, there is a weight function $\rho$ on $\mathbb{F}_q[x_1, \ldots, x_s]/I_q$.

In that case, we can use the following theorem to obtain a lower bound on the minimum distance of affine variety codes $C(I, L)$ for certain subspaces $L$ of $\mathbb{F}_q[x_1, \ldots, x_s]/I_q$.

Our notation is as follows. (Some of this notation was introduced in the remark at the end of section 1.1.) For $R$ having an order function $\rho$, let $\varphi : R \rightarrow \mathbb{F}_q^n$ be a surjective linear map of $\mathbb{F}_q$-algebras. Let $L_k$ be the vector space with $\{f_1, \ldots, f_k\}$ as a basis, where $\{f_i : i \in \mathbb{N}\}$ is a basis of $R$ over $\mathbb{F}_q$ such that $\rho(f_i) < \rho(f_{i+1})$ for all $i \in \mathbb{N}$. Let $E_k = \varphi(L_k)$ and let $C_k$ be the dual code of $E_k$.

Theorem 3.3.9 [HLP, p. 30, Theorem 3.21] Let $\rho$ be a weight function. Then the minimum distance of $E_k$ is at least $n - \rho(f_k)$. If $\rho(f_k) < n$, then $\dim(E_k) = k$.

Proof. See [HLP, p. 30].

Theorem 3.3.10 For an affine variety code $C(I, L)$, where $I$ satisfies the conditions of Proposition 3.3.8, if a set $\{f_1, \ldots, f_m\}$ of representatives (preimages) of the polynomials $\bar{f}_1, \ldots, \bar{f}_m$ spanning $L$ are linearly independent in $\mathbb{F}_q[x_1, \ldots, x_s]/I$ and can be extended to a basis

$$\{f_1, \ldots, f_m, f_{m+1}, \ldots\}$$

of $\mathbb{F}_q[x_1, \ldots, x_s]/I$ so that $\rho(f_i) < \rho(f_j)$ for all $i, j$, $1 \leq i < j$, then the minimum distance of $C$ is at least $n - \rho(f_m)$.

Proof. Since $I$ satisfies the conditions of Proposition 3.3.8, there is a weight
function \( \rho \) on \( R = \mathbb{F}_q[x_1, \ldots, x_s]/I \). By the other assumptions in the statement of the theorem, we have all the conditions for Theorem 3.3.9. It follows that the evaluation code \( E_k \) has minimum distance at least \( n - \rho(f_k) \), by Theorem 3.3.9. But \( E_m = C(I, L) \) by the remark immediately before Section 1.2. Therefore, the minimum distance of \( C \) is at least \( n - \rho(f_m) \). \( \square \)

There is also a lower bound for the minimum distance of certain affine variety codes \( C^{+}(I, L) \). To understand this bound, we need some additional definitions and results.

Recall we defined \( L_k \) above as the vector space with \( \{f_1, \ldots, f_k\} \) as a basis, where \( \{f_i : i \in \mathbb{N}\} \) is a basis of \( R \) over \( \mathbb{F}_q \) such that \( \rho(f_i) < \rho(f_{i+1}) \) for all \( i \in \mathbb{N} \). Recall also that the number \( l(i, j) \) is defined as the smallest positive integer \( k \) such that \( f_i f_j \in L_k \).

**Remark** The function \( l(i, j) \) is strictly increasing in both arguments. This happens because \( \rho \) is an order function and hence satisfies condition (0.3) of Definition 3.3.1. That condition requires that if \( \rho(f) < \rho(g) \) and \( h \neq 0 \), then \( \rho(fh) < \rho(gh) \).

**Definition 3.3.11** [HLP, p. 31] Define \( N_k = \{(i, j) \in \mathbb{N}^2 : l(i, j) = k + 1 \} \). Let \( \nu_k \) be the number of elements of \( N_k \).

**Definition 3.3.12** [HLP, p. 32] The order bounds \( \delta_{ORD}(k) \) and \( \delta_{ORD, \varphi}(k) \) are defined by

\[
\delta_{ORD}(k) = \min\{\nu_{k'} : k' > k\},
\]

\[
\delta_{ORD, \varphi}(k) = \min\{\nu_{k'} : k' \geq k, C_{k'} \neq C_{k'+1}\}.
\]

**Theorem 3.3.13** [HLP, p. 33] The numbers \( \delta_{ORD}(k) \) and \( \delta_{ORD, \varphi}(k) \) are lower bounds for the minimum distance of \( C_k \). Furthermore, \( d_{\min}(C_k) \geq \delta_{ORD, \varphi}(k) \geq \delta_{ORD}(k) \).
**Proof.** See [HLP, p. 33].

**Theorem 3.3.14** For an affine variety code \( C = C^\perp(I, L) \), where \( I \) satisfies the conditions of Proposition 3.3.8, if a set of representatives (preimages) \( \{f_1, \ldots, f_m\} \) for the polynomials \( \tilde{f}_1, \ldots, \tilde{f}_m \) spanning \( L \) are linearly independent in \( \mathbf{F}_q[x, \ldots, x_s]/I \) and can be extended to a basis \( \{f_1, \ldots, f_m, f_{m+1}, \ldots\} \) for \( \mathbf{F}_q[x, \ldots, x_s]/I \) so that \( \rho(f_i) < \rho(f_j) \) for all \( i, j, 1 \leq i < j \), then the minimum distance of \( C \) satisfies \( d_{\min}(C) \geq \delta_{ORD,\varphi}(m) \geq \delta_{ORD}(m) \).

**Proof.** Since \( I \) satisfies the conditions of Proposition 3.3.8, there is a weight function \( \rho \) on \( R = \mathbf{F}_q[x, \ldots, x_s]/I \). Then \( \rho \) is an order function, by Definition 3.3.2. With the other assumptions in the statement of the theorem, we therefore have all the conditions to apply Theorem 3.3.13. Hence, there are lower bounds \( \delta_{ORD,\varphi}(k) \) and \( \delta_{ORD}(k) \) for the minimum distance \( d_{\min} \) of \( C_k \) satisfying \( d_{\min}(C_k) \geq \delta_{ORD,\varphi}(k) \geq \delta_{ORD}(k) \), by Theorem 3.3.13. But, by the remark immediately before Section 1.2, \( C = C^\perp(I, L) = C_m \). Therefore, \( d_{\min}(C) \geq \delta_{ORD,\varphi}(m) \geq \delta_{ORD}(m) \).

**Remark** Feng and Rao [FR] show that, in special cases, for a proper subspace \( L' \) of \( L \), the minimum distance of \( C = C^\perp(I, L') \) may be at least as large as \( \delta_{ORD}(m) \), where the notation is the same as the notation of Theorem 3.3.14. In this case we improve the dimension of the code while retaining a guaranteed minimum distance.

**Example 3.3.15** [FR, p. 1685, Example 3.3]. Let \( I = (x^3y + y^3 + x) \subseteq \mathbf{F}_8[x, y] \). Then \( V(I_8) \) is the affine Klein quartic curve \( x^3y + y^3 + x = 0 \) over \( \mathbf{F}_8 \); there are 22 points in \( V(I_8) \).

Consider the code \( C = C^\perp(I, L) \), where \( L = \langle 1, y, xy, x^2y \rangle \). In [FR], this code is discussed at length and a lower bound computed for its minimum distance, using a method that extends the polynomials \( 1, y, xy, x^2y \) to a basis

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
for $\mathbb{F}_8[x, y]/I_8$ and then computes $\nu(k)$, for $k = 1, \ldots, 22$, (which Feng and Rao call the $N$ sequence). Feng and Rao obtain a lower bound of 3 for the minimum distance of this code.

Now that we have a lower bound on the minimum distance of $C$, we can use the method of Section 3.2 to find $C$'s true minimum distance. First, we calculate the integer $N_0(3)$, for $n = 22$ (since there are 22 points in $V(I_q)$) and $q = 8$.

The partitions of 3 are
\[
\begin{align*}
(3), \\
(2, 1), \\
(1, 1, 1) = (1^3).
\end{align*}
\]
Since $z_1 = 0$, partitions containing a 1 contribute nothing to the sum $N_0$. Thus in this case we need only calculate the term corresponding to the partition 3 of 3.

For $\lambda = (3)$, we have
\[
(z_{\lambda_1})^{r_1} \cdots (z_{\lambda_n})^{r_n} \binom{n}{k} (k! r_1! \cdots r_n!) (d! (\lambda_1 r_1) \cdots (\lambda_n r_n) = (z_3) \binom{22}{1} (3! 3!).
\]
Since
\[
z_3 = (q - 1)^2 - z_2 = (q - 1)^2 - (q - 1)^1 + 0 = 7^2 - 7 = 42,
\]
we obtain
\[
N_0(3) = 42 \cdot 22 = 924.
\]

Next, we calculate the total number of solutions $N(3)$ to the system $S_3$, using the software package Gb [Fau] that computes a Gröbner basis for $S_3$ and also computes the number of points in $V((S_3))$. We then compare $N$ and $N_0$ to see if the code $C$ has codewords of weight 3 or less.

The results of the computation are that the number of solutions to the system $S_3$ is $N(3) = 2100$. Since $N_0(3) = 924$, the difference $N(3) - N_0(3) = 1176$. 
2100 - 924 = 1176 > 0 tells us that there are codewords of weight less than or equal to 3. Since we already know from the lower bound we found earlier that there are no codewords of weight less than 3, we know that all the solutions not counted in \( N_0 \) belong to codewords of weight 3. Consequently, 294 is the number of all the permutations of the nonzero coefficients of the codewords of weight 3. Hence, there are \((N - N_0)/3! = 1176/6 = 196\) codewords of weight 3.
References


Vita

Jeanne Fitzgerald was born on June 20, 1951, in Huntington, West Virginia, and is a citizen of the United States. She received a bachelor of arts degree in German from Virginia Polytechnic Institute in 1972. After working as a computer systems programmer, she earned a master of science degree from Virginia Commonwealth University in 1992. She is currently a candidate for the doctoral degree in mathematics, which will be awarded in August, 1996.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Jeanne Fitzgerald

Major Field: Mathematics

Title of Dissertation: Applications of Gröbner Bases to Linear Codes

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

April 25, 1996