# Computational Algebraic Geometry Applied to Invariant Theory

Ryan Michael Shifler

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mathematics

Ezra A. Brown, Chair
Edward L. Green
Joseph A. Ball

April 30, 2013
Blacksburg, Virginia

Keywords: Groebner Basis, Invariant Theory, Algorithm

# Computational Algebraic Geometry Applied to Invariant Theory

Ryan Michael Shifler

(ABSTRACT)

Commutative algebra finds its roots in invariant theory and the connection is drawn from a modern standpoint. The Hilbert Basis Theorem and the Nullstellenstatz were considered lemmas for classical invariant theory. The Groebner basis is a modern tool used and is implemented with the computer algebra system *Mathematica*. Number 14 of Hilbert's 23 problems is discussed along with the notion of invariance under a group action of $GL_n(\mathbb{C})$. Computational difficulties are also discussed in reference to Groebner bases and Invariant theory.The straitening law is presented from a Groebner basis point of view and is motivated as being a key piece of machinery in proving First Fundamental Theorem of Invariant Theory.

*To Stephen Berstler, Jr.* (September 7, 1989-December 30, 2011)

*...a roommate, a teammate, and a friend.*

# Contents

# List of Figures

# List of Tables

## 0.1 Introduction

The audience of the thesis is directed toward students who have successfully passed a senior level abstract algebra course with minimal coding experience. The thesis is mostly self-contained (i.e. all the the theory needed is either known in a previous course or stated usually with proof).

An ideal $I$ contained in $\mathbb{C}[x_1, \cdots, x_n]$ is finitely generated. It is algorithmically possible to a find Groebner basis $G$ for $I$ which is a computationally preferable generating set of $I$. This allows us to algorithmically decide whether an arbitrary polynomial in $\mathbb{C}[x_1, \cdots, x_n]$ is an element of $I$. $G$ can be used to find the roots of the generators of $I$. A Groebner basis in the single variable polynomial ring case is the greatest common divisors of the generators of $I$. Also, a Groebner basis and reduced Groebner basis are analogous to echelon form and reduced echelon from, respectively. In addition Buchbergers algorithm is the analog of Gauss-Jordan elimination. An explantation of Buchberger's algorithm will be given and is programmed into virtually all computer algebra systems, in particular *Mathematica*. Applications for Groebner bases exist in robotics, graph theory, and, for our purposes, invariant theory.

For an introduction to invariant theory consider the quadratic polynomial $f(x) = ax^2+bx+c$ over $\mathbb{R}$. Suppose the discriminant $b^2 - 4ac > 0$ giving the equation $f(x) = 0$ two distinct solucions $x_1$ and $x_2$. Let $g : \mathbb{R} \to \mathbb{R}$ be a change of variables, that is $g$ is a one-to-one and onto map. Let us consider the solutions to $f(g(x)) = 0$. Since there must exist distinct $x_1', x_2' \in \mathbb{R}$ where $g(x_1') = x_1$ and $g(x_2') = x_2$. That is $f(g(x))$ has exactly two distinct solutions. Similarly if $b^2 - 4ac = 0$ we have $f(x) = 0$ and $f(g(x)) = 0$ have exactly one solution, of multiplicity 2, and if $b^2-4ac < 0$ then we have $f(x) = 0$ and $f(g(x)) = 0$ have no solutions. The point is the sign of the discriminant and the number of solutions remain the same under a change of coordinates [9]. Also note that the actual values of the discriminant and the roots are probably different. Another example of the importance of invariance is in hyperbolic geometry. We need a metric which does not change with an analytic change of variables. It can be shown that

$$\frac{|g'(z)dz|}{1-|g(z)|^2} = \frac{|dz|}{1-|z|^2}$$

for any analytic self-conformal map $g$ of the unit disk which gives rise to the desired metric [8].

Another notion of invariance presented is finding the Groebner bases of ideals where $\sigma(f) \in I$ for all $f \in I$ and for all $\sigma$ in some subgroup of $S_n$. The action, of course, being the permutation of indices of the indeterminates of $f$. One issue that occurs in finding a Groebner basis of an ideal with the aforementioned property is that the symmetry can be lost and sometimes the Groebner basis cannot be solved or is complicated. Examples will be presented.

Invariant theory and commutative algebra have been tied together from the start of the earliest research in the fields. Two well known theorems the Hilbert Basis Theorem and the Nullstellensatz in commutative algebra were lemmas for David Hilbert's work in invariant theory. In the 1960's Bruno Buchberger provided an algorithm to find a Groebner basis which is an application to commutative algebra. The major part of the thesis is going full circle and having commutative algebra solve problems in invariant theory.

The main question which will be attacked is a case of number 14 of Hilbert's 23 problems. The question asked is the ring of invariants of an algebraic group acting on a polynomial ring always finitely generated? The answer is, in general, no with a counterexample provided by Masayoshi Nagata in 1959. In the case presented, however, the answer is yes and our interest is an algorithm to determine the generators.

The final topic studied will be a Groebner basis approach to the Straightening Algorithm. The Straightening Algorithm is a key tool used to prove the First Fundamental Theorem of Invariant Theory. Letting $SL_n(\mathbb{C})$, the set of all $n \times n$ matrices where the determinant is one, act on $\mathbb{C}[x_{i,j}]$ we are able to find the finite generating set. Moreover, this topic will put the notion of a Groebner basis in a more general setting along with a more general notion of Monomial Theory.

Throughout the thesis examples of *Mathematica* code will be given to emphasis the power of the *GroebnerBasis* command. The code is given at the basis level where knowledge of *For*, *Do*, and *While* loops may be the hardest aspect as far as coding is concerned. The literature I found on the material seems to be light on actual implementation and as a result, this thesis attempts to cross that bridge.

Groebner bases, Invariant theory, and computer implementation on *mathematica* are important in physical application, on the one hand with Groebner basis, but also from a pure stand point with invariant theory. An interesting focus, during the time of the original invariant theory research, was the notion of constructivism. For example, the Hilbert Basis Theorem is merely an existence statement which some mathematicians, like Leopold Kronecker, thought was inadequate.

The intertwined topics of invariant theory and commutative algebra-in particular computational algebraic geometry-will be examined throughout the thesis as the topics are fully developed and explained with examples. As a result the reader will have a new perspective of the connections between the topics studied with the motivations being purely mathematical.

# Chapter 1

# Groebner Bases

## 1.1 Ideals and Varieties

A subset of $K[x_1 \cdots x_n]$ is an **ideal** if (i) $0 \in I$, (ii) if $f, g \in I$ then $f + g \in I$, and (iii) if $f \in I$ and $h \in K[x_1 \cdots x_n]$ then $hf \in I$.

The ideal generated by a set of polynomials is the algebraic structure that is used to find the solution set, known as the affine variety. The affine variety is a geometric structure that describes the solutions to all of our polynomial equation. Let $K$ be a field and consider the set, which is an ideal by contruction,

$$I = \langle f_1, f_2, \cdots, f_s \rangle = \left\{ \sum_{i=1}^{s} h_i f_i : h_i \in K[x_1 x_2 x_3 \cdots x_n] \right\}$$

The affine variety of $I$ is

$$V(I) = \{(a_1, \cdots, a_n) \in K^n : f_i(a_1, \cdots, a_n) = 0 \text{ for all } 0 \leq i \leq s\}$$

The following is a theorem which connects the notions of ideals and varieties from [2].

**Theorem:** If $\langle g_1, g_2, \cdots, g_t \rangle = \langle f_1, f_2, \cdots, f_s \rangle$ then $V(g_1, g_2, \cdots, g_t) = V(h_1, h_2, \cdots, h_s)$.

Proof: Let $\langle g_1, g_2, \cdots, g_t \rangle = \langle h_1, h_2, \cdots, h_s \rangle$. Let $a \in V(g_1, g_2, \cdots, g_t)$ so $0 = g_i(a)$ for all $1 \leq i \leq t$. Now each $f_j = \sum_{i=1}^{s} h_{ji} g_i$, for $1 \leq j \leq s$, since $f_j \in \langle g_1, g_2, \cdots, g_s \rangle$. Then $f_j(a) = \sum_{i=1}^{s} h_{j_i}(a) g_i(a) = 0$. Therefore, $a \in V(f_1, f_2, \cdots, f_t)$. By a symmetric argument interchanging $f_j$ and $g_i$, and $s$ and $t$ we prove the reverse inclusion. Therefore, $V(g_1, g_2, \cdots, g_t) = V(h_1, h_2, \cdots, h_s)$.QED.

So, if we have two different generating sets of the same ideal, then the affine variety of the generating sets are the same.

A similar definition for $V(I)$ is as a follows: Let $V \subset K^n$ be an affine variety. Then we set

$$I(V) = \{f \in K[x_1, \cdots, x_n] : f(a_1, \cdots, a_n) = 0 \text{ for all } (a_1, \cdots, a_n) \in V\}$$

## 1.2 Monomial Ordering

This section comes mostly from [2] and [3]. The **multidegree** of a monomial is a vector that is used to describe the exponents in a monomial and is defined as follows. If $x^\alpha = x_0^{r_0} \cdots x_{m-1}^{r_{m-1}}$, then $\alpha = multidegree(x_1^{r_1} \cdots x_n^{r_n}) = (r_1, \cdots, r_n) \in \mathbb{Z}^n_{\geq 0}$

A **monomial order** is a rule used to order the terms in each polynomial. A monomial order is a well ordering, a total ordering, and respects multiplication. An example of what it means to respect multiplication is if $x > y$ then $xz > yz$. Infinitely many monomial orders exist in the multivariable case; lexicographical, graded lexicographical, and graded reverse lex order are the most common.

**Lexicographical Order:** Let $\alpha = (\alpha_1, \cdots, \alpha_n), \beta = (\beta_1, \cdots, \beta_n) \in \mathbb{Z}^n_{\geq 0}$. We say $\alpha >_{lex} \beta$ if, in the vector difference $\alpha - \beta \in \mathbb{Z}^n$, the leftmost nonzero entry is positive.

Lexicographical order is a dictionary order. This works by comparing two monomials by first comparing the exponents of $x_1$. If the exponents are different, then the monomial with the largest exponent corresponding to $x_1$ is the greatest. If the exponents are equal, move on to $x_2$ and repeat the same procedure this will either terminate, or move on to $x_3$. The process is repeated until a difference in exponents is found.

**Graded Lex Order:** Let $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$. We say $\alpha >_{grlex} \beta$ if $|\alpha| = \sum_{i=1}^n \alpha_i > |\beta| = \mu_{i=1}^n \beta_i$, or $|\alpha| = |\beta|$ and $\alpha >_{lex} \beta$.

Graded Lex Order first compares the total degree of the monomial (recall the total degree of $x_1^{r_1} \cdots x_n^{r_n}$ is $r_1 + \cdots + r_n$) and takes the monomial with the largest total degree to be the greatest. If two monomials have the same total degree, then Lexicographical order is used as a tie breaker.

**Graded Reverse Lex Order:** Let $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$. We way that $\alpha >_{grevlex} \beta$ if $|\alpha| = \sum_{i=1}^n \alpha_i > |\beta| = \sum_{i=1}^n \beta_i$, or $|\alpha| = |\beta|$ and the right most nonzero entry for $\alpha - \beta \in \mathbb{Z}^n$ is negative.

Graded Reverse Lex Order first compares total degree. Then, Lex order is used where the variables are reordered so the largest is now the smallest, the second largest is the second smallest, and so on. In other words, lexicographical order uses $x_n > x_{n-1} > \cdots > x_2 > x_1 > x_0$ as the order on the variables.

The leading term of a polynomial $g$, $LT(g)$, is the greatest term in a polynomial under the respective monomial order. Naturally, the leading term will usual depend on the monomial order. For example, let $g = x^3y + xy^2z^2 + x^2y^2z$ with $x > y > z$. Then $x^3y$, $x^2y^2z$, and $xy^2z^2$ are the leading terms under lexicographical, graded lexicographical, and graded reverse lex orders, respectively. From this point forward $x > y > z$ and $x_0 > x_1 > x_2 > \cdots > x_n$ will be the order placed on the variables.

All monomials orders can be described from a set of weight vectors. For an example let $\vec{w_1}, \cdots, \vec{w_s}$ be a set of weight vectors that guarantee a monomial order. Given 2 monomials $x_1^{r_1}x_2^{r_2}x_3^{r_3}\cdots x_n^{r_n}$ and $x_1^{s_1}x_2^{s_2}x_3^{s_3}\cdots x_n^{s_n}$ we must find the multidegree associated with the monomial. So, the $\alpha = multidegree(x_1^{r_1}x_2^{r_2}x_3^{r_3}\cdots x_n^{r_n}) = (r_1, r_2, r_3, \cdots, r_n)$ and $\beta = multidegree(x_1^{s_1}x_2^{s_2}x_3^{s_3}\cdots x_n^{s_n}) = (s_1, s_2, s_3, \cdots, s_n)$. To use the weight vectors we first do $w_1 \cdot \alpha$ and $w_1 \cdot \beta$. If $\vec{w_1} \cdot \alpha > \vec{w_1} \cdot \beta$ then $\alpha > \beta$. If the dot products are equal then do the same computation on $\vec{w_2}$. Repeat this process until a strict inequality is reached moving from $\vec{w_k}$ to $\vec{w_{k+1}}$. Since this is a monomial order the process will terminate [3].

For example lexicographical order can be described with the weight vectors

$$\begin{aligned} \vec{w_1} &= (1,0,0) \\ \vec{w_2} &= (0,1,0) \\ \vec{w_3} &= (0,0,1) \end{aligned}$$

Now consider the monomials $x^3y^2z^2$ and $x^3yz^{100}$. Note $\alpha = multidegree(x^3y^2z^2) = (3,2,2)$ and $\beta = multidegree(x^3yz^{100}) = (3,1,100)$. Then $\alpha \cdot \vec{w_1} = 3 = \beta \cdot \vec{w_1}$ so then we move onto $\vec{w_2}$. Since $\alpha \cdot \vec{w_2} = 2 > 1 = \beta \cdot \vec{w_2}$ we conclude $x^3y^2z^2 > x^3yz^{100}$ under this ordering.

The following is a lemma that proves the monomials $x_1^m, x_2^m, \cdots, x_n^m$ have the same order under every monomial order. The proof uses the fact that all monomial orders can be described with weight vectors.

**Theorem (Example of using weight vectors):** If a monomial order has the restriction $x_1 > x_2 > x_3 > \cdots > x_n$, then $x_1^m > x_2^m > x_3^m > \cdots > x_n^m$.

*Proof:*
Let $\vec{w_1}, \vec{w_2}, \cdots, \vec{w_s}$ be weight vectors that guarantee a monomial order such that $x_1 > x_2 > x_3 > \cdots > x_n$. Let $\alpha = multidegree(x_i)$ and $\beta = multidegree(x_j)$ for $1 \leq i < j \leq n$. Then either $\vec{w_1} \cdot \alpha > \vec{w_1} \cdot \beta$ or $\vec{w_l} \cdot \alpha = \vec{w_l} \cdot \beta$ for $1 \leq l < k \leq s$ and $\vec{w_k} \cdot \alpha > \vec{w_k} \cdot \beta$ since $\alpha > \beta$. Let $\vec{w_l} = (w_{l_1}, \cdots, w_{l_n})$. For the first case we have $w_{1_i} > w_{1_j}$ which implies $w_{1_i} \cdot m > w_{1_j} \cdot m$. For the second case we have $w_{l_i} = w_{l_j}$ for $1 \leq l < k$ and $w_{k_i} > w_{k_j}$. Then, $w_{l_i} \cdot m = w_{l_j} \cdot m$ for $1 \leq l < k$ and $w_{k_i} \cdot m > w_{k_j} \cdot m$. This shows $multidegree(x_i^m) > multidegree(x_j^m)$, or

$x_i^m > x_j^m$.QED

## 1.3 Multivariable Division Algorithm

This section comes from [2]. We move to another important preliminary idea, the multivariable division algorithm. The algorithm will be necessary for many computations and theory in this paper. The algorithm is analogous to the single variable division algorithm learned in high school. Division of

$$f \in K[x_1 x_2 \cdots x_n]$$

by

$$f_1, f_2, \cdots, f_s \in K[x_1 x_2 \cdots x_n]$$

allows for $f$ to written as

$$f = c_1 f_1 + c_2 f_2 + \cdots + c_s f_s + r$$

where the leading term of $r$ is not divisible by the leading term of any of the divisors.

Consider $x^3yz+x^3y^2z^2$ divided by $\{x^2yz + yz, xy + xyz^2\}$ under lexicographical order. Then,

$$x^3yz + x^3y^2z^2 = (x + xyz)(x^2yz + yz) + (-y)(xy + xyz^2) + xy^2 - xyz.$$

Now after seeing what it does we will state the algorithm as presented in [2]. The first step is to fix a monomial order and the proceed as below.

```
Input: $f_1, \cdots, f_s, f$
Output: $a_1, \cdots, a_s, r$
$a_1 := 0, \cdots, a_s := 0; r = 0$
$p := f$
WHILE $p \neq 0$ DO
    $i := 1$
    $divisionoccured := false$
    While $i \leq s$ AND $divisionoccured = false$ DO
        IF $LT(f_i)$ divides $p$ THEN
            $a_i := a_i + LT(p)/LT(f_i)$
            $p := p - (LT(p)/LT(f_i))f_i$
            $divisionoccured := true$
        ELSE
            $i := i + 1$
```

```
IF $divisionoccured = false$ THEN
$r := r + LT(p)$
$p := p - LT(p)$
```

This invokes a new definition. Let $f$ and $f_1, \cdots, f_s$ be as they are in the algorithm. Let $F = \{f_1, \cdots, f_s\}$. Then $\bar{f}^F$ is the remainder output by the algorithm above.

The following is the example above being computed in *Mathematica*. The output of the *PolynomialReduce* command is $\{\{a_1, \cdots, a_s\}, r\}$ based on the notation in the algorithm.

In[1]:= $f = x^3 * y * z + x^3y^2z^2$;
F $= \{x^2 * y * z + y * z, x * y + x * y * z^2\}$;

In[2]:= PolynomialReduce[f, F, {x, y, z}, MonomialOrder − > Lexicographic]

Out[2]= $\{\{x + xyz, -y\}, xy^2 - xyz\}$

This concludes our short study of the multivariable division algorithm.

## 1.4 Hilbert Basis Theorem

We continue this introduction to Groebner basis with some famous results that many texts, including [2], [3], [7], [1], call theorems. However, these famous results in commutative algebra were actually lemmas in David Hilbert's study on invariant theory. We will call these famous results lemmas and we will see their power in our study on invariant theory.

The first is the Hilbert Basis Lemma. This says every ideal of $K[x_1 \cdots x_n]$ is finitely generated. Care needs to be taken when looking at generators of a polynomials ring of an ideal. For example, the ideal $I = \langle x^2, y^3 \rangle = \{h_1x^2 + h_2y^3 : h_1, h_2 \in K[x_1 \cdots x_n]\}$. If we let $h_1 = 0$ and $h_2 = y$, we have $y^4 \in I$. However, $y^4$ can not be written in terms of the indeterminates $x^2$ and $y^3$. So, $I \neq K[x^2, y^3]$.

To present the Hilbert Basis Lemma two lemmas will be presented whose results are relevant to the study of Groebner bases. The first is known as Dickson's Lemma and is a first glimpse at finite generation of an ideal. The proof can also be found in [2].

**Dickson's Lemma:** Let $\left\langle x^\alpha : \alpha \in A \subset \mathbb{Z}^n_{\geq 0} \right\rangle$. Then $I$ can be written in the form $I = \left\langle x^{\alpha(1)}, \cdots, x^{\alpha(s)} \right\rangle$, where $\alpha(1), \cdots, \alpha(s) \in A$. In particular, $I$ has a finite basis.

Proof by induction on the number of variables: Let $n = 1$ and $I = \langle x^\alpha | \alpha \in A \subset \mathbb{Z} \rangle$. Choose $\beta \leq \alpha$ for all $\alpha \in A$. Thus, $I = \left\langle x^\beta \right\rangle$.

Let $n > 1$ and suppose the theorem holds for $n-1$. We will work in $K[x_1, \cdots, x_{n-1}, y]$ so every monomial has the form $x^\alpha y^p$ where $\alpha \in \mathbb{Z}^{n-1}_{\geq 0}$ and $p \in \mathbb{Z}$.

Let $I = \left\langle x^\alpha y^p : (\alpha, p) \in A \subset \mathbb{Z}^n_{\geq 0} \right\rangle \subset K[x_1, \cdots, x_{n-1}, y]$ be a monomial ideal.

Let $J = \langle x^\alpha : x^\alpha y^m \in I \text{ for some } m \in \mathbb{Z}_{\geq 0} \rangle$. Our inductive hypothesis holds so $J = \left\langle x^{\alpha(1)}, \cdots, x^{\alpha(s)} \right\rangle$. For each $i$ between 1 and $s$ we have $x^{\alpha(i)} y^{m_i} \in I$ for some $m_i \geq 0$. Let $m = \max\{m_i\}$. Now consider $J_k = \left\langle x^\beta : x^\beta y^k \in I \text{ for some } m \in \mathbb{Z}_{\geq 0} \right\rangle$ for each $k, 0 \leq k \leq m-1$. Then hour theorem holds in this case so $J_k = \left\langle x^{\alpha_k(1)}, \cdots, x^{\alpha_k(s_k)} \right\rangle$. The claim is $I$ is generated by some subset of

$$D = \{x^{\alpha(1)} y^m, \cdots, x^{\alpha(s)} y^m\} \bigcup \left( \bigcup_{k=0}^{m-1} \{x^{\alpha_k(1)} y^k, \cdots, x^{\alpha_k(s_k)} y^k\} \right).$$

Now we claim that every monomial in $I$ is divisible by some element in $D$. Let $x^\alpha y^p \in I$. If $p \geq m$ then $x^\alpha y^p$ is divisible by some $x^{\alpha(i)} y^m$ by the construction of $J$. If $p \leq m-1$ then $x^\alpha y^p$ is divisible by some element of $J_p$ by construction. Thus we have proven the ever monomial in $I$ is divisible by some element in $D$ thus $\langle D \rangle = I$.

Now we know that $I = \left\langle x^{\delta_1}, \cdots, x^{\delta_j} \right\rangle$ where $x^{\delta_i} \in D$. (Note the difference in notation $\delta \in \mathbb{Z}^n_{\geq 0}$). Now $x^{\delta_i} \in I$ thus $x^{\delta_i}$ is divisible by some $x^{\alpha(i)} y^{p(i)}$ with $(\alpha(i), p(i)) \in A$. Thus $I = \left\langle x^{\alpha(1)} y^{p(1)}, \cdots, x^{\alpha(j)} y^{p(j)} \right\rangle$. Thus the theorem is proven.QED.

Thus we have proven finite generation for ideals with the form $\left\langle x^\alpha : \alpha \in A \subset \mathbb{Z}^n_{\geq 0} \right\rangle$. We now follow with the next important lemma. The proof can also be found in [2]

**Lemma:** Let $I \subset K[x_1 \cdots x_n]$ be an ideal. Then there exists a set $A$ where $\langle LT(I) \rangle = \left\langle x^\alpha : \alpha \in A \subset \mathbb{Z}^n_{\geq 0} \right\rangle$ and there exist $g_1, \cdots, g_t \in I$ such that $\langle LT(I) \rangle = \langle LT(g_1), \cdots, LT(g_t) \rangle$.

Proof: The leading monomials $LM(g)$ of elements $g \in I - \{0\}$ generate the ideal $\langle LM(g) : g \in I - \{0\} \rangle$. Since $LM(g)$ and $LT(g)$ differ by a nonzero constant, this ideal equals $\langle LT(g) : g \in I - \{0\} \rangle = \langle LT(I) \rangle$.

Since $\langle LT(I) \rangle$ is generated by the monomials $LM(g)$ for $g \in I - \{0\}$, Dickson's Lemma assures use that $\langle LT(I) \rangle = \langle LM(g_1), \cdots, LM(g_t) \rangle$ for finitely many $g_1, \cdots, g_t \in I$. Since $LM(g_i)$ differs from $LT(g_i)$ by a nonzero constant, it follows that $\langle LT(I) \rangle = \langle LT(g_1), \cdots, LT(g_t) \rangle$. QED.

We now present the Hilbert Basis Lemma. The proof of the following theorem can also be found in [1] and [2].

**Hilbert Basis Lemma:** Every ideal $I \subset K[x_1 \cdots x_n]$ has a finite generating set. That is $I = \langle g_1, \cdots, g_t \rangle$ for some $g_1, \cdots, g_t \in I$.

Proof: If $I = \{0\}$, we take our generating set to be $\{0\}$, which is finite. If $I$ contains some nonzero polynomial, then a generating set $g_1, \cdots, g_t$ for $I$ can be constructed as follows. By the above proposition, there are $g_1, \cdots, g_t \in I$ such that $\langle LT(I) \rangle = \langle LT(g_1), \cdots, LT(g_t) \rangle$. We claim that $I = \langle g_1, \cdots, g_t \rangle$.

First note that $\langle g_1, \cdots, g_t \rangle \subset I$ since $g_1, \cdots, g_t \in I$. Let $f \in I$ be any polynomial. If we apply the division algorithm to divide $f$ by $\{g_1, \cdots, g_t\}$, then we get an expression of the form $f = a_1 g_1 + \cdots + a_t g_t + r$ where no term of $r$ is divisible by any of $LT(g_1), \cdots, LT(g_t)$. We claim that $r = 0$. To see this, note that $r = f - a_1 g_1 - \cdots - a_t g_t \in I$. If $r \neq 0$, then $LT(r) \in \langle LT(I) \rangle = \langle LT(g_1), \cdots, LT(g_t) \rangle$. By the proof of Dickson's lemma it follows that $LT(r)$ must be divisible by some $LT(g_i)$. This contradicts what it means to be a remainder, consequently r must be zero. Thus $f \in I$.QED.

## 1.5 Groebner basis

Everything in this section is found from [2]. Given a set of polynomials, $g_1, g_2, \cdots, g_q$, finding the variety of $I = \langle g_1, g_2, \cdots, g_q \rangle$ can be computationally difficult for certain generators. This leads to the notion of a Groebner Basis.

**Groebner basis:** Fix a monomial order $>$.A finite set of polynomials $\{h_1, \cdots, h_a\} \subset I$ is a *Groebner basis* if $\langle LT(h_1), \cdots, LT(h_a) \rangle = \langle LT(I) \rangle$ where $\langle LT(I) \rangle$ is the ideal generated by all the leading terms in the ideal $I$.

**Theorem:** (i)Every ideal $I \subset K[x_1 \cdots x_n]$ has a Groebner basis $G$ under any monomial order and (ii)$I = \langle G \rangle$.

Proof: (i) is an immediate result of the Proposition immediately proceeding Hilbert's Basis (Theorem) Lemma. (ii) is by the construction of the Hilbert Basis (Theorem) Lemma's proof.QED.

A Groebner basis for an ideal $I$ can be arrived at algorithmically. A Groebner basis for an ideal $I$ is a set of polynomials that is easier to work with in computational setting with few exceptions. Groebner bases are usually dependent on the monomial order. For example, consider $I = \langle x - z^4, y - z^5 \rangle$. The Groebner bases under lexicographical and graded lexico-

graphic order are respectively $\{x - z^4, y - z^5\}$ and
$\{xz - y, z^4 - x, yz^3 - x^2, y^2z^2 - x^3, x^4 - y^3z\}$.

One example of a Groebner basis is a linear set of equations in echelon form. Another example is in the single variable case. Let $f, g \in K[x]$ where $K$ is a field. Then, $\{gcd(f, g)\}$ is a Groebner basis of $\langle f, g\rangle$.

A **reduced Groebner basis** for an ideal $I \subset K[x_1 \cdots x_n]$ is a Groebner basis $G$ for $I$ such that for all distinct $p, q \in G$, no monomial appearing in $p$ is a multiple of $LT(q)$. We follow with another definition which will be of use later. A **monic Groebner basis** is a reduced Groebner basis in which the leading coefficients of every polynomial is 1, or is empty if $I = \langle 0\rangle$[3].

## 1.6 Buchberger's Criterion and Algorithm

Everything in this section can be found in atleast [2]. Buchberger's criterion is satisfied for set of polynomials if and only if the set is a Groebner basis. Buchberger's criterion will be used to be sure that our sets of polynomials is a Groebner basis. The **S-Polynomial** is defined as

$$S(g_i, g_j) = \frac{y^\alpha}{LT(g_i)}g_i - \frac{y^\alpha}{LT(g_j)}g_j$$

where $y^\alpha = LCM(LT(g_i), LT(g_j))$.

**Theorem:** Let $I$ be a polynomial ideal. Then a basis $G = \{g_1, \cdots, g_a\}$ for $I$ is a Groebner basis for $I$ if and only if for all pairs $i \neq j$, the remainder on division of $S(g_i, g_j)$ by $G$ is zero.

Proof: This proof can be found on page 85 of [2].

**Theorem:** Given a finite set $G \subset k[x_1, \cdots, x_n]$, suppose that we have $f, g \in G$ such that the leading monomials of $f$ and $g$ are relatively prime, then we know the remainder on division of $S(f, g)$ by $G$ is zero.

Proof: This proof can be found on page 104 of [2].

Buchberger's Algorithm is let $I = \langle f_1, \cdots, f_s\rangle \neq \{0\}$ be a polynomial ideal. Then a Groebner basis for $I$ can be constructed in the following way:

Input: $F = (f_1, \cdots, f_s)$
Output: a Groebner basis $G = (g_1, \cdots, g_t)$ for $I$, with $F \subset G$

```
$G := F$
REPEAT $G' := G$
FOR each pair $\{p, q\}$, $p \neq q$ in $G'$ DO
$S := S(\bar{p}, q)^{G}$
IF $S \neq 0$ THEN $G := G \cup \{S\}$
UNTIL $G = G'$.
```

The following is an example of computing a Groebner basis in the computer algebra system *Mathematica*. First we give two examples for ways to compute reduced Groebner bases for both, Lexicographical and graded reverse lexicographic order. The first uses the built in mathematical monomial order, and the second uses weight matrices. The rows in the weight matrices are the weight vectors discussed above with $w_1$ being row one, $w_2$ being row two and so on.

$In[1] := F = \{x - z^4, y - z^4\};$
$V = \{x, y, z\};$

In[2]:= GroebnerBasis[F, V, MonomialOrder $- >$ Lexicographic]

$Out[2] = \{y - z^4, x - z^4\}$

In[3]:= M = {{1, 0, 0}, {0, 1, 0}, {0, 0, 1}};

In[4]:= GroebnerBasis[F, V, MonomialOrder $- >$ M]

$Out[4] = \{y - z^4, x - z^4\}$

In[5]:= GroebnerBasis[F, V, MonomialOrder $- >$ DegreeReverseLexicographic]

$Out[5] = \{-x + y, -y + z^4\}$

In[6]:= M1 = {{1, 1, 1}, {0, 0, -1}, {0, -1, 0}};

In[7]:= GroebnerBasis[F, V, MonomialOrder $- >$ M1]

$Out[7] = \{-x + y, -y + z^4\}$

Recall that we have assumed $x > y > z$ as the order on the variables, the Groebner Basis command returned a reduced Groebner bases, and Groebner bases are dependent on monomial orders. Finally, a result that may be found in [2] states a reduced Groebner basis for a given monomial order is unique up to scalar multiplication.

# Chapter 2

# Symmetric Ideals with an Original Result

## 2.1 Definitions and Examples of Desired Result

Everything in this section and the next is from [12]. We discuss how invariants and Groebner bases are related beginning with a result we will spend the next few pages proving. In this section we will consider $K[x_0 \cdots x_{m-1}]$ where $K$ is a field. The results examples and results are true for any monomial order with $x_0 > x_1 > \cdots > x_{m-1}$ by using the Theorem on page 5. Let $S_m$ act on $K[x_0 \cdots x_{m-1}]$ in the natural way. Let $\sigma \in S_m$ then we say an ideal $I \subset K[x_0 \cdots x_{m-1}]$ is $\sigma$**-Symmetric** if $\sigma(I) = I$. For example $I = \langle x_0^2 + x_1^2 + x_2^2, x_0 x_1 x_2 \rangle$ is $(0, 1, 2)$-Symmetric.

We will move on to the specific case when $\sigma = (0, 1, \cdots, m-1)$ and the ideal is generated by the orbits of the polynomial $x_0^n + x_d^n$, that is $I_d^m = \langle \sigma(x_0^n + x_d^n) : \sigma \in \langle (0, 1, \cdots, m-1) \rangle \rangle$. From the construction $I_d^m$ is a $(0, 1, 2, \cdots, m-1)$-Symmetric ideal. Now define an identical ideal but in a different way using modular arithmetic. We will see this is beneficial when trying to prove our big result for this section.

Let $1 \leq d \leq \frac{m}{2}$ be the distance between variables where $d = d(x_i, x_j) = |j - i|$ or $d = m - d(x_i, x_j)$. Let the ideal generated by a circulant system of polynomials be

$$F_d^m = \left\langle x_i^n - x_{d_i}^n : 0 \leq i \leq m-1, d_i \equiv i + d (\text{mod m}) \right\rangle$$

Nothing is lost restricting $d$ to be between 1 and $\frac{m}{2}$. If $d > \frac{m}{2}$, then using $m - d$ will lead to the same results. For example, the $m = 6$ and $d = 4$ case is

$$\begin{array}{rcllllll}
0 & = & -x_0^n & & +x_2^n & & & \\
0 & = & & -x_1^n & & +x_3^n & & \\
0 & = & & & -x_2^n & & +x_4^n & \\
0 & = & & & & -x_3^n & & +x_5^n \\
0 & = & x_0^n & & & & -x_4^n & \\
0 & = & & x_1^n & & & & -x_5^n
\end{array}$$

Each polynomial equation in this system is a polynomial in the $m = 6$ and $d = 2$ case multiplied by $-1$. So the $d = 2$ and $d = 4$ cases for $m = 6$ results in the same system.

It is not difficult to see that $I_d^m = F_d^m$ for $1 \leq d \leq \frac{m}{2}$.

**Example 1:**
Continuing with the circulant system of polynomials previously discussed

$$\begin{array}{rcllllll}
0 & = & x_0^n & & -x_2^n & & & \\
0 & = & & x_1^n & & -x_3^n & & \\
0 & = & & & x_2^n & & -x_4^n & \\
0 & = & & & & x_3^n & & -x_5^n \\
0 & = & -x_0^n & & & & +x_4^n & \\
0 & = & & -x_1^n & & & & +x_5^n
\end{array}$$

In this case $m = 6$ and $d = 2$ and the ideal we are working with is

$$\begin{array}{rcl}
F & = & \langle x_0^n - x_2^n, x_1^n - x_3^n, x_2^n - x_4^n, x_3^n - x_5^n, -x_0^n + x_4^n, -x_1^n + x_5^n \rangle \, . \\
\text{Let } F' & = & \{x_0^n - x_2^n, x_1^n - x_3^n, x_2^n - x_4^n, x_3^n - x_5^n, -x_0^n + x_4^n, -x_1^n + x_5^n\}
\end{array}$$

We will now check to see if $F'$ is a Groebner basis. As already stated if two polynomials have relatively prime leading terms, the division of the corresponding S-polynomial by $F'$ results in a remainder of 0. Hence, it is only necessary to check polynomials with the same leading terms. Notice how the first two and last two equations in our list share leading terms while the rest are in the middle. Inorder to generalize this method for any $d$ and $m$, let $A = \{x_0^n - x_2^n, x_1^n - x_3^n\}$, $B = \{x_2^n - x_4^n, x_3^n - x_5^n\}$, and $C = \{x_0^n - x_4^n, x_1^n - x_5^n\}$ so that $F = \langle A \cup B \cup C \rangle$. So now we must only check the polynomials in $A$ and $C$. To finish we will find the two S-polynomials and divide them by $F'$ and check to see if the remainder is 0. First,

$$S(x_0^n - x_2^n, x_0^n - x_4^n) \quad = \quad -x_2^n + x_4^n$$

and $-x_2^n + x_4^n = -1(x_2^n - x_4^n) + 0$. As for our second two polynomials,

$$S(x_1^n - x_3^n, x_1^n - x_5^n) \quad = \quad -x_3^n + x_5^n$$

and $-x_3^n + x_5^n = -1(x_3^n - x_5^n) + 0$. So we have a remainder of 0 for both and hence we have a Groebner basis. The division of the S-polynomial by $F'$ only used polynomials in $B$, and as it turns out, the size of $B$ is what forces $F'$ to be a Groebner basis. A **Universal Groebner basis** is a Groebner basis under every monomial ideal. So, $F$ is a universal Groebner basis. Then, we can find a reduced universal Groebner basis for $F$. This is a simple computation in this case and we find $\{x_0^n - x_4^n, x_1^n - x_5^n, x_2^n - x_4^n, x_3^n - x_5^n\}$.

**Example 2:**
Let $m = 7$ and $d = 2$ for $F$. We claim the set of generators of $F$ is not a Groebner basis. We will use a similar construction to a more general proof like we did in the previous example. The corresponding polynomial equations are

$$\begin{array}{rcccccccc}
0 & = & x_0^n & & -x_2^n & & & & \\
0 & = & & x_1^n & & -x_3^n & & & \\
0 & = & & & x_2^n & & -x_4^n & & \\
0 & = & & & & x_3^n & & -x_5^n & \\
0 & = & & & & & x_4^n & & -x_6^n \\
0 & = & -x_0^n & & & & & +x_5^n & \\
0 & = & & -x_1^n & & & & & +x_6^n
\end{array}$$

The ideal we are concerned with is

$$\begin{aligned}
F &= \langle x_0^n - x_2^n, x_1^n - x_3^n, x_2^n - x_4^n, x_3^n - x_5^n, x_4^n - x_6^n, -x_0^n + x_5^n, -x_1^n + x_6^n \rangle . \\
\text{Let } F' &= \{x_0^n - x_2^n, x_1^n - x_3^n, x_2^n - x_4^n, x_3^n - x_5^n, x_4^n - x_6^n, -x_0^n + x_5^n, -x_1^n + x_6^n\}
\end{aligned}$$

Like we did early, we will divide up $F'$ into three sets like we did in the previous example. Let $A = \{x_0^n - x_2^n, x_1^n - x_3^n\}$, $B = \{x_2^n - x_4^n, x_3^n - x_5^n, x_4^n - x_6^n\}$, $C = \{x_0^n - x_5^n, x_1^n - x_6^n\}$. So now we have $F = \langle A \cup B \cup C \rangle$. We must check to see that the S-polynomials of two polynomials in $F$ does not result in 0. Lets choose the first equation in $A$ and $B$.

$$S(x_0^n - x_2^n, x_0^n - x_5^n) = -x_2^n + x_5^n$$

Proceed using the multivariable division algorithm we find

$$\begin{aligned}
-x_2^n + x_5^n &= -1(x_2^n - x_4^n) - 1(x_4^n - x_5^n) \\
&= -1(x_2^n - x_4^n) - 1(x_4^n - x_6^n) - 1(x_6^n - x_5^n)
\end{aligned}$$

Now we can see that the remainder is $x_5^n - x_6^n$ since no polynomial in $F'$ has $x_5^n$ as a leading term. So now we can conclude that $F'$, when $m = 7$ and $d = 2$, is not a Groebner basis.

The following lemmas and proposition show $F'$ is a universal Groebner basis if and only if $d$ divides $m$. Redefining the generators of $F$ into 3 disjoint sets will allow the proofs to be completed. Let $A = \{x_i^n - x_{i+d}^n : 0 \leq i \leq d-1\}$, $B = \{x_i^n - x_{i+d}^n : d \leq i \leq m-d-1\}$, and $C = \left\{x_i^n - x_{i+(m-d)}^n : 0 \leq i \leq d-1\right\}$. Now we have $F = \langle A \cup B \cup C\rangle$ and $F' = A{\cup}B{\cup}C$. In order to prove this final result we must distinguish between $m < 3d$ and $m \geq 3d$. Proposition 1 and 2 will cover the case of when $m < 3d$ and are relatively straightforward computations.

## 2.2 Main Results

**Proposition 1:** Let $m < 3d$. If $m = 2d + r$ where $1 \leq r \leq d-1$, then $F'$ is not a Groebner Basis.

*Proof:*
Proceeding with checking Buchberger's criterion, consider $S(x_0^n - x_d^n, x_0^n - x_{m-d}^n) = x_d^n - x_{m-d}^n$. Since $m = 2d + r$ where $1 \leq r \leq d-1$ we know $x_d^n - x_{2d}^n \in B$. So by the division algorithm we have

$$x_d^n - x_{m-d}^n = (x_d^n - x_{2d}^n) + (x_{2d}^n - x_{m-d}^n).$$

Since $m < 3d$ implies $m - d < 2d$, the leading term of $x_{2d}^n - x_{m-d}^n$ is $x_{m-d}^n$. Since $x_{m-d}^n - x_{2d}^n$ is a none zero remainder in $F'$ since $x_{m-d}^n$ is not the leading term of any polynomial in $F'$. So $F'$ is not a Groebner Basis.Q.E.D.

**Propostion 2:** If $m = 2d$, then $F'$ is a universal Groebner basis.

*Proof:*
First note $\left\langle x_i^n - x_{d_i}^n : 0 \leq i \leq m-1, d_i \equiv i + d (\text{mod m})\right\rangle = \langle A \bigcup C\rangle$ since $m = 2d$ so $B$ is empty. Also, we see each polynomial in $A$ is the same as one in $C$ and vice versa. So, we can eliminate each polynomial in $C$ since they are redundant and consider $\langle A\rangle$. Since each leading term is disjoint from the other so $A$ is a Groebner basis. Then, $F'$ is a Groebner basis.Q.E.D.

So now we know that if $m < 3d$ then $F$ is a universal Groebner basis if and only if $m = 2d$.

For the $m \geq 3d$ case we start of with two preliminary lemmas that will be used to be sure that each step in the multivariable division algorithm can be made and that the algorithm terminates.

**Lemma 2:** Let $m = (q+2)d$ for some positive integer $q$, equivalently $d$ divides the order of $B$, and $0 \leq i \leq d-1$. If $1 \leq w \leq q$ then $i + wd < i + (m-d)$ and $d \leq i + wd \leq m - d - 1$.

*Proof:*

Let $1 \leq w \leq q$ then, $i + wd \leq i + qd < i + (q+1)d = i + (q+2)d - d = i + m - d$ and $d \leq i + wd \leq i + qd = i + m - 2d \leq m - d - 1$ for $0 \leq i \leq d - 1$.

**Lemma 3:** If $m = (q+2)d + r$ for $1 \leq r \leq d-1$, then $d \leq kd \leq m - d - 1$ for $1 \leq k \leq q+1$ and $m - d < (q+2)d$.

*Proof:*
Let $1 \leq k \leq q+1$ then, $d \leq kd \leq (q+1)d < (q+1)d+1 \leq (q+1)d+r = (q+2)d+r-d = m-d$. Also, $m - d = (q+1)d + r < (q+2)d$.

Lemma 2 is used in the forward direction of Proposition 3 and Lemma 3 is used for the converse. The converse is proved by the contrapositive.

**Proposition 3:** Let $m \geq 3d$. $m = (q+2)d$ for some positive integer $q$ if and only if $F'$ is a university Groebner basis.

*Proof:*
First, Lemma 1 shows there is only one way to order the terms in each polynomial so we need to show that $F'$ is a Groebner Basis. Also, since division of the S-polynomial by $F'$ of polynomials with relatively prime leading terms results in a remainder of zero, it suffices to check polynomials with the same leading term.

$S(x_i^n - x_{i+d}^n, x_i^n - x_{i+(m-d)}^n) = x_{i+d}^n - x_{i+(m-d)}^n$ for $0 \leq i \leq d-1$.

Let $R_k = x_{i+kd}^n - x_{i+(m-d)}^n$, so in the first iteration of the multivariable division algorithm we have

$$x_{i+d}^n - x_{i+(m-d)}^n \quad = \quad 1 * (x_{i+d}^n - x_{i+2d}^n) + (x_{i+2d}^n - x_{i+(m-d)}^n)$$

We can see the previous equations takes on the form $R_1 = 1 * f + R_2$ for some $f \in B$.

Does $R_k = 1 * g + R_{k+1}$ for some $g \in B$ provided $k \leq q$? We have

$$x_{i+kd}^n - x_{i+(m-d)}^n \quad = \quad 1 * (x_{i+kd}^n - x_{i+(k+1)d}^n) + (x_{i+(k+1)d}^n - x_{i+(m-d)}^n)$$

According to Lemma 2, $(x_{i+kd}^n - x_{i+(k+1)d}^n) \in B$ since $k \leq q$. So now in our multivariable division algorithm we have a sequence of "remainders" $R_1, R_2, \cdots, R_s$. The following shows that $R_{q+1} = 0$,

$$x_{i+qd}^n - x_{i+(m-d)}^n \quad = \quad 1 * (x_{i+qd}^n - x_{i+(q+1)d}^n)$$

Since $m - d = (q+2)d - d = (q+1)d$, so $R_q \in B$. So now we can conclude that $F$ is a universal Groebner basis.

For the converse let $m = (q+2)+r$ where $q$ is positive integer and $1 \leq r < d-1$. Consider the division of $S(x_0^n - x_d^n, x_0^n - x_{m-d}^n) = x_d^n - x_{m-d}^n$ by $F'$. Then proceeding with the first iteration of the multivariable division algorithm we have

$$x_d^n - x_{m-d}^n = (x_d^n - x_{2d}^n) + (x_{2d}^n - x_{m-d}^n).$$

We know from Lemma 3, $x_d^n - x_{2d}^n \in B$. Now consider $x_{kd}^n - x_{m-d}^n$ for $1 \leq k \leq q+1$. Using the division algorithm we have

$$x_{kd}^n - x_{m-d}^n = (x_{kd}^n - x_{(k+1)d}^n) + (x_{(k+1)d}^n - x_{m-d}^n).$$

We know $(x_{kd}^n - x_{(k+1)d}^n) \in B$ by Lemma 3. We will eventually need to divide $x_{(q+1)d}^n - x_{m-d}^n$. So by the division algorithm that we have,

$$x_{(q+1)d}^n - x_{m-d}^n = (x_{(q+1)d}^n - x_{(q+2)d}^n) + (x_{(q+2)d}^n - x_{m-d}^n).$$

Once again by Lemma 3 we know $x_{(q+1)d}^n - x_{(q+2)d}^n \in B$. Also, we have $x_{m-d}^n - x_{(q+2)d}^n$ as the remainder since the leading term is $x_{m-d}^n$ from Lemma 3 and no other polynomial in $F'$ shares the same leading term. Hence, $F'$ is not a Groebner basis. We have now shown the following:

**Theorem:** The generators of $F$, which is denoted by $F'$, is a universal Groebner basis if and only if $d$ divides $m$.

For an example for finding the reduced Groebner basis, let's consider $F$ for $m = 9$ and $d = 3$. The polynomials are listed below.

$$\begin{array}{ccccccccc}
x_0^n & & & -x_3^n & & & & & \\
& x_1^n & & & -x_4^n & & & & \\
& & x_2^n & & & -x_5^n & & & \\
& & & x_3^n & & & -x_6^n & & \\
& & & & x_4^n & & & -x_7^n & \\
& & & & & x_5^n & & & -x_8^n \\
-x_0^n & & & & & & +x_6^n & & \\
& -x_1^n & & & & & & +x_7^n & \\
& & -x_2^n & & & & & & +x_8^n
\end{array}$$

We know from the previous theorem we have a universal Groebner basis. We will now construct a reduced universal Groebner basis. Since the last 3 polynomials share a leading term with the first three polynomials, we can remove them and be left with

$$\begin{array}{ccccccccc}
x_0^n & & & -x_3^n & & & & & \\
& x_1^n & & & -x_4^n & & & & \\
& & x_2^n & & & -x_5^n & & & \\
& & & x_3^n & & & -x_6^n & & \\
& & & & x_4^n & & & -x_7^n & \\
& & & & & x_5^n & & & -x_8^n
\end{array}$$

No monomial in any polynomial can be the leading term of another polynomial. So $x_0^n$, $x_1^n$, $x_2^n$, $x_3^n$, $x_4^n$, $x_5^n$ may only exist in this system as leading terms. So to remove $x_3^n$ from the first polynomial, we will add $(x_0^n - x_3^n) + (x_3^n - x_6^n) = x_0^n - x_6^n$ to the system and remove $x_0^n - x_3^n$. So the Groebner basis is now

$$\begin{array}{ccccccccc}
x_0^n & & & & & & -x_6^n & & \\
 & x_1^n & & & -x_4^n & & & & \\
 & & x_2^n & & & -x_5^n & & & \\
 & & & x_3^n & & & -x_6^n & & \\
 & & & & x_4^n & & & -x_7^n & \\
 & & & & & x_5^n & & & -x_8^n
\end{array}$$

To remove $x_4^n$ and $x_5^n$ from their respective polynomials are similar process will be used. So we will add $(x_1^n - x_4^n) + (x_4^n - x_n^7) = x_1^n - x_7^n$ and $(x_2^n - x_5^n) + (x_5^n - x_n^8) = x_2^n - x_8^n$ and remove $x_1^n - x_4^n$ and $x_2^n - x_5^n$. So the system is now

$$\begin{array}{ccccccccc}
x_0^n & & & & & & -x_6^n & & \\
 & x_1^n & & & & & & -x_7^n & \\
 & & x_2^n & & & & & & -x_8^n \\
 & & & x_3^n & & & -x_6^n & & \\
 & & & & x_4^n & & & -x_7^n & \\
 & & & & & x_5^n & & & -x_8^n
\end{array}$$

All the conditions necessary to be a reduced Universal Groebner basis are satisfied. The corollary builds off this example and the proof is parallel to the example above.

**Corollary:** If $m = (q+2)d$ for some positive integer $q$, then $F$ has a reduced Universal Groebner basis(RUGB) taking the form $\left\{x_{l+kd}^n - x_{l+(m-d)}^n : 0 \le l \le d-1, 0 \le k \le \frac{m}{d} - 2\right\}$.

*Proof:*
From the theorem we know $F'$ is a universal Groebner basis. Also, since the polynomials in $C$ have the same leading terms as those in $A$, we can eliminate each polynomial in $C$. So now we have $F = \left\langle x_i^n - x_{i+d}^n : 0 \le i \le m-d-1 \right\rangle$. Define $f_l = x_l^n - x_{l+d}^n$ for simplicity. We will now construct a reduced Groebner basis.

$$f_{l+kd} + f_{l+(k+1)d} + f_{l+(k+2)d} + \cdots + f_{l+qd} \;=\; x_{l+kd}^n - x_{l+(q+1)d}^n = x_{l+kd}^n - x_{l+(m-d)}^n$$

for $0 \le l \le d-1$ and $0 \le k \le \frac{|B|}{d} = \frac{m}{d} - 2$. So we now have

$$\left\{x_{0+kd}^n - x_{0+(m-d)}^n, x_{1+kd}^n - x_{1+(m-d)}^n, x_{2+kd}^n - x_{2+(m-d)}^n, \cdots, x_{(d-1)+kd}^n - x_{m-1}^n : 0 \le k \le \frac{m}{d} - 2\right\}$$

and we must check to be sure it is infact a Reduced Groebner Basis. With the use of the following inequality,

$$x_0^n < x_1^n < \cdots < x_{(d-1)+(\frac{m}{d}-2)d}^n = x_{m-d-1}^n < x_{m-d}^n < x_{1+(m-d)}^n \cdots < x_{m-1}^n$$

we can see that the criteria necessary for a Reduced Groebner Basis has been satisfied.Q.E.D.

## 2.3 Symmetric Ideals and Computational Difficulties

See [11] for more information on what this section is presenting. To a more general and interesting connection between Groebner bases and $(0, 1, \cdots, m-1)$-Symmetric ideals we consider the ideal generated by the following polynomials:

$$\begin{aligned} x_0 + \cdots x_{m-1} \\ x_0x_1 + x_1x_2 + \cdots + x_{m-2}x_{m-1} + x_{m-1}x_0 \\ \vdots \\ x_0x_1 \cdots x_{m-2} + x_1x_2 \cdots x_{m-1} + \cdots + x_{m-2}x_{m-1} \cdots x_{m-4} + x_{m-1}x_0 \cdots x_{m-3} \\ x_0 \cdots x_{m-1} - 1. \end{aligned}$$

This ideal is called *cyclic*$(n)$ and the variety $V(cyclic(m))$ are called cyclic $n$-roots. Finding a Groebner bases for $m$ greater than 7 is computationally difficult and Stefan Steidel makes use of the fact *cyclic*$(n)$ is a $(0, 1, \cdots, m-1)$-Symmetric ideal in [S] to compute a Groebner basis for $m = 9$.

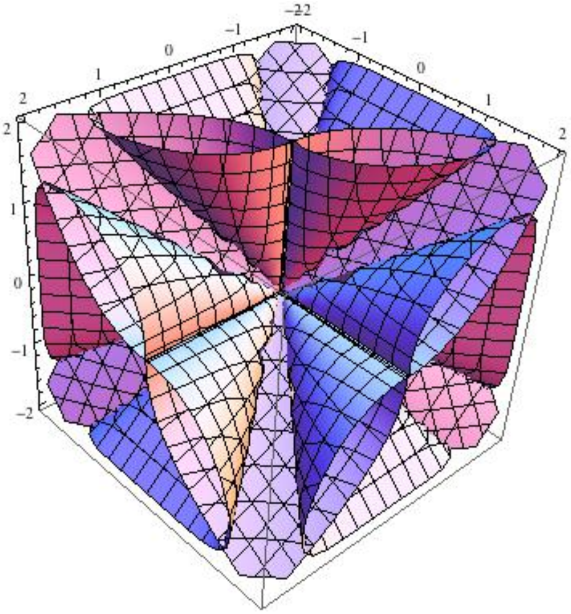
There are other systems that carry on this notion of circulation that are much more elaborate. One such example is

$$\begin{array}{rcl} 0 & = & y^nx^m \quad +x^ny^m \quad +z^nz^m \\ 0 & = & z^nx^m \quad +y^ny^m \quad +x^nz^m \\ 0 & = & x^nx^m \quad +z^ny^m \quad +y^nz^m \end{array}$$

Notice in this system $x^m, y^m$, and $z^m$ are fixed while $y^n, x^n$, and $z^n$ are shifted.

The following is an image of the graphs of each of the three polynomials above with $n = 3$ and $m = 4$.

Figure 2.1: Graph for $n = 3$ and $m = 4$ case



The variety of these polynomials is the intersection of these three graphs. The circulation in the original polynomial system seems to have a connection with the graphical symmetries seen above. The following is the list of the polynomials that compose the reduced Groebner basis under graded lexicographical order.

$y^7 + x^4z^3 + x^3z^4, x^4y^3 + x^3y^4 + z^7, x^7 + y^4z^3 + y^3z^4, x^2y^7z - xy^7z^2 + 2y^7z^3 + y^6z^4 + y^4z^6 + x^2yz^7 - xy^2z^7 + 2y^3z^7, -y^{10} + x^3y^4z^3 - x^3y^3z^4 + z^{10}, x^3y^7 + y^7z^3 + y^6z^4 - x^3z^7 + x^2yz^7 - xy^2z^7 + y^3z^7, xy^{10} + y^{11} - xz^{10} - z^{11}, xy^7z^3 + y^8z^3 + xy^6z^4 + 2y^7z^4 + x^3z^8, -y^8z^3 - y^7z^4 + xy^4z^6 + x^3yz^7 - x^2y^2z^7 + 2xy^3z^7 + y^4z^7 + y^3z^8, xy^8z^3 + y^9z^3 + xy^7z^4 + 3y^8z^4 + y^7z^5 - xy^4z^7 + x^2y^2z^8 - 2xy^3z^8 - y^4z^8 - y^3z^9, x^2y^4z^6 + xy^5z^6 + x^2y^3z^7 + 3xy^4z^7 + y^5z^7 + xy^3z^8 + y^4z^8, xy^8z^3 + x^2y^6z^4 + 4xy^7z^4 + y^8z^4 + xy^6z^5 - y^7z^5 - y^6z^6 - y^4z^8 - x^2yz^9 + xy^2z^9 - 2y^3z^9, 5y^{13} + 7y^{12}z + 7y^{11}z^2 + 9y^{10}z^3 + 6y^9z^4 + 2xy^7z^5 + 2y^8z^5 + y^7z^6 - 2xy^5z^7 - y^6z^7 - 2y^5z^8 - 6y^4z^9 - 9y^3z^{10} - 7y^2z^{11} - 7yz^{12} - 5z^{13}, -27y^{13} - 39y^{12}z - 39y^{11}z^2 - 57y^{10}z^3 + 12xy^8z^4 - 40y^9z^4 - 6y^8z^5 + 6xy^6z^6 - 3y^7z^6 + 26xy^5z^7 + 9y^6z^7 + 18xy^4z^8 + 20y^5z^8 + 2xy^3z^9 + 32y^4z^9 + 45y^3z^{10} + 39y^2z^{11} + 39yz^{12} + 29z^{13}, ?9y^{13} + 13y^{12}z + 13y^{11}z^2 + 4xy^9z^3 + 27y^{10}z^3 + 28y^9z^4 + 6y^8z^5 - 6xy^6z^6 + y^7z^6 - 26xy^5z^7 - 7y^6z^7 - 30xy^4z^8 - 20y^5z^8 - 6xy^3z^9 - 20y^4z^9 - 15y^3z^{10} - 13y^2z^{11} - 13yz^{12} - 11z^{13}, 3y^{12}z + 3y^{11}z^2 + 3y^{10}z^3 + y^9z^4 + 3xy^6z^6 + 3y^7z^6 + 4xy^5z^7 + 9y^6z^7 + 3xy^4z^8 + 4y^5z^8 + xy^3z^9 + y^4z^9 + 3x^2z^{11} - 3xyz^{11} - 3yz^{12} - 2z^{13}, -9y^{13} - 15y^{12}z - 15y^{11}z^2 - 21y^{10}z^3 - 16y^9z^4 + 6xy^6z^6 + 15y^7z^6 + 8xy^5z^7 + 9y^6z^7 + 6xy^4z^8 + 8y^5z^8 + 2xy^3z^9 + 20y^4z^9 + 6x^2yz^{10} - 6xy^2z^{10} + 27y^3z^{10} + 15y^2z^{11} + 15yz^{12} + 11z^{13}, 48y^{13}z + 73y^{12}z^2 + 74y^{11}z^3 + 23y^{10}z^4 + 6y^9z^5 + 5y^8z^6 - 5y^6z^8 - 6y^5z^9 - 23y^4z^{10} - 74y^3z^{11} - 73y^2z^{12} - 48yz^{13}, 24y^{14} - 19y^{12}z^2 - 14y^{11}z^3 + 43y^{10}z^4 + 30y^9z^5 + y^8z^6 - y^6z^8 - 30y^5z^9 - 43y^4z^{10} + 14y^3z^{11} + 19y^2z^{12} - 24z^{14}, -115y^{12}z^2 + 1858y^{11}z^3 + 6355y^{10}z^4 - 642y^9z^5 + 1033y^8z^6 - 48y^7z^7 + 935y^6z^8 + 3618y^5z^9 +$

$2717y^4z^{10} + 542y^3z^{11} + 115y^2z^{12} + 4512xz^{13} + 2112yz^{13} + 10416z^{14}, 1577y^{12}z^2 + 730y^{11}z^3 - 8873y^{10}z^4 - 4026y^9z^5 + 469y^8z^6 - 48y^7z^7 + 1499y^6z^8 + 7002y^5z^9 + 4512xy^3z^{10} + 17945y^4z^{10} + 6182y^3z^{11} - 1577y^2z^{12} + 2112yz^{13} + 5904z^{14}, -167y^{12}z^2 - 166y^{11}z^3 + 1447y^{10}z^4 + 454y^9z^5 - 187y^8z^6 + 48y^7z^7 - 277y^6z^8 + 1504xy^4z^9 - 422y^5z^9 - 1495y^4z^{10} - 730y^3z^{11} + 167y^2z^{12} - 608yz^{13} - 1392z^{14}, 91y^{12}z^2 + 158y^{11}z^3 - 1939y^{10}z^4 - 1542y^9z^5 - 121y^8z^6 - 168y^7z^7 + 1128xy^5z^8 + 241y^6z^8 + 678y^5z^9 + 979y^4z^{10} + 346y^3z^{11} - 91y^2z^{12} + 624yz^{13} + 1488z^{14}, -51y^{12}z^2 - 510y^{11}z^3 + 2099y^{10}z^4 + 2174y^9z^5 + 105y^8z^6 + 1504xy^6z^7 + 528y^7z^7 + 807y^6z^8 - 318y^5z^9 - 1123y^4z^{10} - 322y^3z^{11} + 51y^2z^{12} - 672yz^{13} - 1776z^{14}, 1197y^9z^6 - 2459y^8z^7 - 7490y^7z^8 - 29252y^6z^9 - 53363y^5z^{10} - 33881y^4z^{11} - 9260y^3z^{12} - 5198y^2z^{13} - 19649yz^{14} - 23705z^{15}, 63y^{10}z^5 + 130y^8z^7 + 364y^7z^8 + 1426y^6z^9 + 2587y^5z^{10} + 1612y^4z^{11} + 439y^3z^{12} + 253y^2z^{13} + 931yz^{14} + 1105z^{15}, 1197y^{11}z^4 - 1867y^8z^7 - 5047y^7z^8 - 20287y^6z^9 - 37273y^5z^{10} - 22945y^4z^{11} - 6196y^3z^{12} - 4042y^2z^{13} - 13993yz^{14} - 15502z^{15}, 399y^{12}z^3 - 27y^8z^7 - 574y^7z^8 - 1607y^6z^9 - 3051y^5z^{10} - 2341y^4z^{11} - 767y^3z^{12} + 198y^2z^{13} + 35yz^{14} - 905z^{15}, 717y^6z^{10} + 2615y^5z^{11} + 1997y^4z^{12} + 488y^3z^{13} - 320y^2z^{14} + 298yz^{15} + 1090z^16, 717y^7z^9 - 5191y^5z^{11} - 4282y^4z^{12} + 1352y^3z^{13} + 5731y^2z^{14} + 4105yz^{15} - 812z^{16}, 717y^8z^8 + 2546y^5z^{11} + 1238y^4z^{12} - 7657y^3z^{13} - 15008y^2z^{14} - 12983yz^{15} - 4088z^{16}, y^3z^{14} + 6y^2z^{15} + 6yz^{16} + 5z^{17}, y^4z^{13} - yz^{16}, y^5z^{12} - 7y^2z^{15} - 6yz^{16} - 6z^{17}, y^2z^{16} + yz^{17} + z^{18}$

Groebner bases are usually computationally preferable, however the reduced Groebner basis is large. The complexity is of interest since the original polynomial system is clean and has nice graphical symmetries.

# Chapter 3

# More Computational Algebraic Geometry

## 3.1 Elimination Theory

Everything from this section can be found in [2]. Other useful references are [1] and [3]. Thus far we have discussed a what a Groebner basis is, but why do we care. One easy reason is the ideal member problem. Let $I \subset K[x_1 \cdots x_n]$ be an ideal. Let $f \in K[x_1 \cdots x_n]$. Is $f \in I$? To answer this question we compute a Groebner basis $G$ of $I$ and find $\bar{f}^G$. Then the theorem below answers the question.

**Theorem:** Let $\{g_1, \cdots, g_t\}$ be a Groebner basis for an ideal $I \subset K[x_1, \cdots, x_n]$ and let $f \in K[x_1, \cdots, x_n]$. Then there is a unique $r \in K[x_1, \cdots, x_n]$ with the property that there is a $g \in I$ with $f = g + r$.

Proof: Using the division algorithm we have $f = a_1g_1 + \cdots + a_tg_t + r$ where $a_1, \cdots, a_t \in K[x_1, \cdots, x_n]$. Then let $g = a_1g_1 + \cdots + a_tg_t \in I$. So $f = g + r$.

For the uniqueness of $r$ let $f = g + r = g' + r'$ where $g, r, g', r'$ are found using the division algorithm. Then $r - r' = g - g' \in I$. If $r \neq r'$ then $LT(r - r') \in LT(I)$. So $LT(g_i)|LT(r - r')$. for some $g_i$. But this is a contradiction so $r = r'$.QED.

So we can infer from above that $f \in I$ if and only if $\overline{f}^G = 0$, since $r$ was found by using the division algorithm.

Another important reason to study Groebner bases is finding solution sets, the varieties, corresponding to a system of polynomials. For an easy example, let's consider the system in

$\mathbb{C}[xyz]$

$$x^2 + z = 0$$
$$y^2 - z = 0$$
$$z^2 - 1 = 0.$$

One can readily check that that $\{x^2 + z, y^2 - z, z^2 - 1\}$ is a reduced Groebner basis under lexicographic order with $x > y > z$ since the $GCD$ of the leading terms for any pair is 1. In addition neither $x^2, y^2, z^2$ divide $z$.

Now we will consider $\{x^2 + z, y^2 - z, z^2 - 1\} \cap \mathbb{C}[z] = \{z^2 - 1\}$. So we now have $z^2 - 1 = 0$ if $z = \pm 1$. Then for $\{x^2 + z, y^2 - z, z^2 - 1\} \cap \mathbb{C}[yz] = \{y^2 - z, z^2 - 1\}$ we have $(1, 1), (1, 1), (i, -1), (-i, -1)$ as the solution set for $y^2 - z = 0$ and $z^2 - 1 = 0$. Finally by repeating the step above one more time we find that $V(x^2 + z, y^2 - z, z^2 - 1) = \{(i, 1, 1), (-i, 1, 1), (i, -1, 1), (-i, -1, 1), (1, i, -1), (1, -i, -1), (-1, i, -1), (-1, -i, -1)\}$.

The procedure seen above does generalize with a much deeper theory known as elimination theory. We begin with a definition given $I = \langle f_1, \cdots, f_s \rangle \subset K[x_1 \cdots x_n]$ the $lth$ **elimination ideal** $I_l$ is the ideal of $K[x_{l+1} \cdots x_n]$ defined by $I_l = I \cap K[x_{l+1} \cdots x_n]$. For a check, $I$ and $K[x_{l+1} \cdots x_n]$ are subrings of $K[x_1 \cdots x_n]$ so $I_l$ is a ring. Let $f \in I_l$ and $h \in K[x_{l+1} \cdots x_n]$. Since $f \in I$ we may say that $f \cdot h \in I$ since $h \in K[x_{l+1} \cdots x_n] \subset K[x_1 \cdots x_n]$. Also $f \cdot h \in K[x_{l+1} \cdots x_n]$ since $K[x_{l+1} \cdots x_n]$ is a ring. So $I_l$ is an ideal in $K[x_{l+1} \cdots x_n]$.

We follow with a theorem that gives the relationship between Groebner bases and elimination ideals. This will be a first glimpse to understand the general strategy of the example given above. This result can also be found in [2].

**The Elimination Theorem:** Let $I \subset K[x_1 \cdots x_n]$ be an ideal, let $0 \leq l \leq n$ and let $G$ be a Groebner basis of $I$ with respect to a monomial ordering where any monomial involving $x_1, \cdots, x_l$ is greater than all monomials in $K[x_{l+1} \cdots x_n]$. Then the set $G_l = G \cap K[x_{l+1} \cdots x_n]$ is a Groebner basis of the $l$-th elimination ideal $I_l$. When using lexicographic order with $x_1 > x_2 > \cdots x_n$ the theorem is true for all $0 \leq l \leq n$.

Proof: Let $0 \leq l \leq n$ and note that $G_l = G \cap K[x_{l+1} \cdots x_n] \subset I \cap K[x_{l+1} \cdots x_n] = I_l$ since $I \subset G$. We need to show that $\langle LT(I_l) \rangle = \langle LT(G_l) \rangle$ to satisfy the definition of a Groebner basis. Let $f \in \langle LT(G_l) \rangle$. So $LT(g_i)$ divides $f$ for some $g_i \in G_l$. Since $g_i \in I_l$ as well, we see that $f \in \langle LT(I_l) \rangle$, thus proving the easy inclusion $\langle LT(G_l) \rangle \subset \langle LT(l_l) \rangle$.

Let $f \in \langle LT(l_i) \rangle$, that is $f = LT(f')$ for some $f' \in I_l$. Now $f' \in I$ which indicates $LT(g)$ divided $LT(f')$ for some $g \in G$. Since $f' \in I_l$, this means that $LT(G)$ involves only the variables $x_{l+1}, \cdots, x_n$. Since we are using an order in which any monomial involving $x_1, \cdots, x_l$ is greater than all monomials in $K[x_{l+1} \cdots x_n]$, $LT(g) \in K[x_{l+1} \cdots x_n]$ implies $g \in K[x_{l+1} \cdots x_n]$. Thus we may say that $g \in G_l$. So we have proven the desired equality.

QED

## 3.2 Extension Theorem

Relating this theorem back to the previous example we see that $\{z^2-1\}$ and $\{y^2-z, z^2-1\}$ are Groebner bases of $\langle x^2+z, y^2-z, z^2-1\rangle \cap K[z]$ and $\langle x^2+z, y^2-z, z^2-1\rangle \cap K[yz]$, respectively. There is another theorem, the Extension Theorem [2], that is used as a second part to elimination theory and will be stated and not proved. The geometric version of the Extension Theorem will then be stated and not proved so we can then prove Hilbert's Nullstellensatz, David Hilbert's Lemma in invariant theory that is taken as a theorem in commutative algebra.

**The Extension Theorem:** Let $\bar{K}$ be an algebraically closed field. Let $I = \langle f_1, \cdots, f_s\rangle \subset \bar{K}[x_1 \cdots x_n]$ and let $I_1$ be the first elimination ideal of $I$. For each $1 \leq i \leq s$, write $f_i$ in the form $f_i = g_i(x_2, \cdots, x_n)x_1^{N_i}$ + terms in which $x_1$ has degree $< N_i$, where $N_i > 0$ and $g_i \in \mathbb{C}[x_2 \cdots x_n]$ is nonzero. Suppose that we have a partial solution $(a_2, \cdots, a_n) \in V(I_1)$. If $(a_2, \cdots, a_n) \notin V(g_1, \cdots, g_s)$, then there exists $a_1 \in \mathbb{C}$ such that $(a_1, \cdots, a_n) \in V(I)$.

With this theorem we can justify the steps we took to find the solutions in the example above by first finding the solution to $z^2-1=0$, $y^2-z=z^2-1=0$, and finally $x^2+z=y^2-z=z^2-1=0$.

**Geometric Extension Theorem:** Give $V = V(f_1, \cdots, f_s) \subset \bar{K}^n$, let $g_i$ be as the Extension Theorem. If $I_1$ is the first elimination ideal of $\langle f_1, \cdots, f_s\rangle$, then we have the equality in $\bar{K}^{n-1}$

$$V(I_l) = \pi_1(V) \cup (V(g_1, \cdots, g_s) \cap V(l_1)),$$

where $\pi_1 : \bar{K}^n \to \bar{K}^{n-1}$ is a projection onto the last $n-1$ components.

## 3.3 Nullstellensatz

We start with the statement and proof of the Weak Nullstellensatz which will then be used to prove the Hilbert Nullstellensatz.

**Weak Nullstellensatz:** Let $\bar{K}$ be an algebraically closed field and let $I \subset \bar{K}[x_1 \cdots x_n]$

be an ideal satisfying $V(I) = \emptyset$. Then $I = \bar{K}[x_1 \cdots x_n]$.

Proof by induction: If $n = 1$ and $I \subset \bar{K}[x]$ satisfies $V(I) = \emptyset$ which means $I = \langle c \rangle$ where $c$ is a nonzero constant since $\bar{K}[x]$ is a P.I.D and $\bar{k}$ is algebraically closed. Then $1 = c \cdots (1/c) \in I$. So $I = \bar{K}[x]$.

Assume the result has been proved for the polynomial ring in $n-1$ variable, which we write as $\bar{K}[x_2 \cdots x_n]$. Consider any ideal $I = \langle f_1, \cdots, f_s \rangle \subset \bar{K}[x_1 \cdots x_n]$ for which $V(I) = \emptyset$. We may assume that $f_1$ is not a constant since, otherwise, there is nothing to prove. So, suppose $f_1$ has total degree $N \geq 1$. We will next change coordinates so that $f_1$ has an especially nice form. Namely, consider the linear change of coordinates

$$\begin{aligned} x_1 &= \widetilde{x_1}, \\ x_2 &= \widetilde{x_2} + a_2\widetilde{x_1}, \\ &\vdots \\ x_n &= \widetilde{x_n} + a_n\widetilde{x_1}. \end{aligned}$$

where $a_i$ are as-yet-to-be-determined constants in $\bar{K}$. Substitute for $x_1, \cdots, x_n$ so that $f_1$ has the form

$$\begin{aligned} f_1(x_1, \cdots, x_n) &= f_1(\widetilde{x_1}, \widetilde{x_2} + a_2\widetilde{x_1}, \cdots, \widetilde{x_n} + a_n\widetilde{x_1}) \\ &= c(a_2, \cdots, a_n)\widetilde{x_1}^N + \text{terms in which } \widetilde{x_1} \text{ has degree} < N. \end{aligned}$$

Since $\bar{K}$ is a field, it does not have zero divisors, or else, $f_1$ does not have degree $N$. So $c(a_2, \cdots, a_n)$ is nonzero for some $a_2, \cdots, a_n$.

With this choice of $a_2, \cdots, a_n$, under the coordinate change above every polynomial $f \in \bar{K}[x_1 \cdots x_n]$ goes over to a polynomial $\widetilde{f} \in \bar{K}[\widetilde{x_1} \cdots \widetilde{x_n}]$. Note that we still have $V(\widetilde{I}) = \emptyset$ since if the transformed equations had solutions, so would the original ones. Furthermore, if we can show that $1 \in \widetilde{I}$, then $1 \in I$ will follow since constants are unaffected by the tilde operation.

By the previous paragraph $f_1 \in \widetilde{I}$ transforms to $\widetilde{f_1} \in \widetilde{I}$ with the property

$$\widetilde{f_1}(\widetilde{x_1}, \cdots, \widetilde{x_n}) = c(a_2, \cdots, a_n)\widetilde{x_1}^N + \text{ terms in which } \widetilde{x_1} \text{ has degree} < N,$$

where $c(a_2, \cdots, a_n) \neq 0$. This allows us to use the Geometric Extension Theorem, to relate $V(\widetilde{I})$ with its projection into the subspace $\bar{K}$ with coordinates $\widetilde{x_2}, \cdots, \widetilde{x_n}$. Let $\pi_1 : \bar{K}^n \to \bar{K}^{n-1}$ be the projection mapping onto the last $n-1$ components. If we set $\widetilde{I_1} = \widetilde{I} \cap \bar{K}[\widetilde{x_2} \cdots \widetilde{x_n}]$ as usual, then parital solutions in $\bar{K}^{n-1}$ always extend. By the induction hypothesis, it follows that $I_1 = \bar{K}[\widetilde{x_2} \cdots \widetilde{x_n}]$. But this implies $1 \in \widetilde{I_1} \subset \widetilde{I}$, and this completes the proof.QED.

**Hilbert Nullstellensatz:** Let $\bar{K}$ be an algebraically closed field. If $f, f_1, \cdots, f_s \in \bar{K}[x_1 \cdots x_n]$ are such that $f \in I(V(f_1, \cdots, f_s))$, then there exits an integer $m \geq 1$ such that $f^m \in \langle f_1, \cdots, f_s \rangle$(and conversely).

Proof: Given a nonzero polynomial $f$ which vanishes at every common zero of the polynomial $f_1, \cdots, f_s$, we must show that there exists an integer $m \geq 1$ and polynomials $A_1, \cdots, A_s$ such that $f^m = \sum_{i=1}^{s} A_i f_i$.

Consider the ideal $\widetilde{I} = \langle f_1, \cdots, f_s, 1 - yf \rangle \subset \bar{K}[x_1 \cdots x_n y]$, where $f, f_1, \cdots, f_s$ are as above. We claim that $V(\widetilde{I}) = \emptyset$. To see this, let $(a_1, \cdots, a_n, a_{n+1}) \in \bar{K}^{n+1}$. Either (i) $(a_1, \cdots, a_n)$ is a common zero of $f_1, \cdots, f_s$, or (ii) $(a_1, \cdots, a_n)$ is not a common zero of $f_1, \cdots, f_s$.

In case (i) $f(a_1, \cdots, a_n) = 0$ since $f$ vanishes at any common zero of $f_1, \cdots, f_s$. Thus the polynomial $1-yf$ takes the value $1-a_{n+1}f(a_1, \cdots, a_n) = 1 \neq 0$ at the point $(a_1, \cdots, a_n, a_{n+1})$. In particular $(a_1, \cdots, a_n, a_{n+1}) \notin V(\widetilde{I})$.

In case (ii), for some $i$, $1 \leq i \leq s$, we must have $f_i(a_1, \cdots, a_n) \neq 0$. Viewing $f_i$ as a function of $n+1$ variables which does not depend on the last variable, we have $f_i(a_1, \cdots, a_n, a_{n+1}) \neq 0$. In particular, we again conclude that $(a_1, \cdots, a_n, a_{n+1}) \notin V(\widetilde{I})$. Since $(a_1, \cdots, a_n, a_{n+1}) \in \bar{K}^{n+1}$ was arbitrary, we conclude that $V(\widetilde{I}) = \emptyset$ as claimed.

Now apply the Weak Nullstellensatz to conclude that $1 \in \widetilde{I}$. That is,

$$1 = \sum_{i=1}^{s} p_i(x_1, \cdots, x_n, y) f_i + q(x_1, \cdots, x_n, y)(1 - yf)$$

for some polynomials $p_i, q \in \bar{K}[x_1 \cdots x_n y]$. Now set $y = 1/f(x_1, \cdots, x_n)$. Then relation above implies that

$$1 = \sum_{i=1}^{s} p_i(x_1, \cdots, x_n, 1/f) f_i.$$

Multiply both sides of this equation by a power $f^m$, where $m$ is chosen sufficiently large to clear all the denominators. This yields $f^m = \sum_{i=1}^{s} A_i f_i$, for some polynomials $A_i \in \bar{K}[x_1 \cdots x_n]$. QED.

Radical ideals and their relationship with Groebner bases have an important application to invariant theory. An ideal is a $I$ is **radical** if $f^m \in I$ for some integer $m \geq 1$ implies that $f \in I$. As a consequence let $V$ be a variety and $f^m \in I(V)$. If $x \in V$ then $f^m(x) = 0$ if and only if $f(x) = 0$, so $f \in I(V)$. So $I(V)$ is a radical ideal [2].

We will denote and define the **radical ideal** of and ideal $I \in K[x_1 \cdots x_n]$ by

$$\sqrt{I} = \{f : f^m \in I \text{ for some integer } m \geq 1\}$$

**Theorem:** Let $I \in K[x_1, \cdots, x_n]$ be an ideal, then $\sqrt{I}$ is an ideal and $I \subset \sqrt{I}$.

Proof: First $0 \in I$, thus $0 \in \sqrt{I}$ by definition.

Let $f, g \in \sqrt{I}$. So, $f^i, g^j \in I$ for some $i, j \in \mathbb{Z}^+$. Then every term in the expansion $(f+g)^{i+j}$ has the form $f^m g^n$. Either $m \geq i$ or $n \geq j$ since $m < i$, and $n < j$ implies $m + n < i + j$ since we should have $m + n = i + j$. Thus, either $f^m$ or $g^n$ is in $I$ so each term is in $I$. Thus $(f+g)^{i+j} \in I$. We may conclude $f + g \in \sqrt{I}$.

Let $f \in \sqrt{I}$ and $h \in K[x_1 \cdots x_n]$. So $f^i \in I$ for some $i \in \mathbb{Z}^+$ and $h^i \in K[x_1 \cdots x_n]$. Thus $(fh)^i = f^i h^i \in I$. So, $fh \in \sqrt{I}$. Therefore $\sqrt{I}$ is an ideal.

Let $f \in I$, then $f^1 \in I$ which implies $f \in \sqrt{I}$. So $I \subset \sqrt{I}$.QED.

We have developed enough machinery to present our next major theorem.

**Nullstellenstaz Theorem:** Let $K$ be an algebraically closed field. If $I \subset K[x_1 \cdots x_n]$ is an ideal then $\mathbf{I}(V(I)) = \sqrt{I}$.

Proof: Let $K$ be an algebraically closed field and $I \subset K[x_1 \cdots x_n]$ be an ideal. Let $f \in \sqrt{I}$. Then $f^i \in I$ for some $i \in \mathbb{Z}^+$. So $f^i$ vanishes on $V(I)$ be definition and as a consequence $f$ vanishes on $V(I)$. So $f \in \mathbf{I}(V(I))$.

Let $f \in \mathbf{I}(V(I))$. Then $f$ vanishes on $V(I)$ be definition. Now, by the Hilbert Nullstellensatz, there exist $i \in \mathbb{Z}^+$ such that $f^i \in I$. So $f \in \sqrt{I}$ by definition.

We have proven both directions of the inclusion so $\mathbf{I}(V(I)) = \sqrt{I}$.QED.

We have now presented and proven the two lemmas necessary for our study of invariant theory. Moreover, notice that we used Hilbert's Basis Lemma to prove theorems about Groebner basis, and then used Groebner basis theory to justify a step in the proof of the Hilbert Nullstellensatz.

# Chapter 4

# Invariant Theory

## 4.1 Invariant Rings

This paper is about invariant theory using Groebner bases as an aid, but first it is important to mention that invariant theory is useful for the study of Groebner basis. Examples can be found in [7] chapter 2.6, and in many papers. One such example is [11] where an alternative algorithm is presented for finding a Groebner basis when certain invariant conditions are satisfied. We now discuss the key ideas of this invariant theory related to this paper.

We are working in the polynomial ring $K[x_1 \cdots x_n]$, and let $\vec{x} = (x_1, \cdots, x_n)$. For notational purposes, for all $f \in K[x_1 \cdots x_n]$ we define $f(\vec{x}) := f(x_1, \cdots, x_n)$. Let $G$ be a group that *acts* on $\{\vec{x}\}$. We will say $K[x_1 \cdots x_n]^G$ is the set of all polynomials $f \in K[x_1 \cdots x_n]$ where $f(\sigma\vec{x}) = f(\vec{x})$ for all $\sigma \in G$.

We will define a polynomial $f$ to be **symmetric** if $f \in K[x_1 \cdots x_n]^{S_n}$ were $f(\sigma\vec{x}) := f(x_{\sigma(1)}, \cdots, x_{\sigma(n)})$ for all $\sigma \in S_n$. Let $GL_n(K)$ denote the general linear group of $n \times n$ matrices with entries coming from $K$. We let $GL_n(K)$ act on $\{\vec{x}\}$ by matrix multiplication as $A \cdot \vec{x}^T$ for all $A \in GL_n(K)$. An immediate question that one can ask is: what are $K[x_1 \cdots x_n]^{S_n}$ and $K[x_1 \cdots x_n]^{GL_n(K)}$? This is one of the main questions in invariant theory. To answer this question satisfactorily, we must find $f_1, \cdots, f_s \in K[x_1 \cdots x_n]$, a finite number $s$, such that $K[x_1 \cdots x_n]^{S_n} = K[f_1 \cdots f_s]$, a similar notion for a finite subgroup of $GL_n(K)$, and for $GL_n(K)$. Part of this notion is a case of Hilbert's 14th problem which asks "Is the ring of invariants of an algebraic group acting on a polynomial always finitely generated?" We will eventually see that the answer to this question is yes for both $S_n$ and finite subgroups of $GL_n(K)$. We wish to find these generators.

We define $\mathbb{S} := \{\sigma_r = \sum_{i_1 < i_2 < \cdots < i_r} x_{i_1} \cdots x_{i_r} \colon\ 1 \le r \le n\}$ to be the set of **elementary symmetric polynomials** and $\mathbb{P} = \{s_k = x_1^k + \cdots + x_n^k : 1 \le k \le n\}$ to be the set of **power**

**sums**. By our construction we have $\mathbb{S}, \mathbb{P} \subset K[x_1 \cdots x_n]^{S_n}$.The answer for $S_n$ is elementary involving $\mathbb{S}$ and $\mathbb{P}$ for two different solutions which will be presented later. The answer is not simple for a finite subgroup $GL_n(K)$ and $GL_n(K)$ in general. Heavier mathematical machinery is needed and these solutions, non-constructive and constructive, will eventually be presented.

## 4.2 Preliminary Results

These results can all be found [2] as either a stated result or problem.

**Theorem:** $K \leq_R K[x_1 \cdots x_n]^G \leq_R K[x_1 \cdots x_n]$.

Proof: (i) First note that $K[x_1 \cdots x_n]^G$ contains $K$ since the group $G$ acts on the indeterminates, thus all constant polynomials must be invariant.

(ii) It suffices to show that $K[x_1 \cdots x_n]^G$ is nonempty, which was shown by part (i), and closed under subtraction and multiplication. Let $f, g \in K[x_1 \cdots x_n]^G$. Let $\sigma \in G$, and $\sigma(x_1, \cdots, x_n) = (x_{i_1}, \cdots, x_{i_n})$ then

$$\begin{aligned}((f-g)(\sigma(x_1,\cdots,x_n))) &= (f-g)(x_{i_1},\cdots,x_{i_n})) \\ &= f(x_{i_1},\cdots,x_{i_n})) - g(x_{i_1},\cdots,x_{i_n})) \\ &= f(x_1,\cdots,x_n) - g(x_1,\cdots,x_n) \\ &= (f-g)(x_1,\cdots,x_n).\end{aligned}$$

Therefore, $f - g \in K[x_1 \cdots x_n]^G$. Also,

$$\begin{aligned}((fg)(\sigma(x_1,\cdots,x_n))) &= (fg)(x_{i_1},\cdots,x_{i_n})) \\ &= f(x_{i_1},\cdots,x_{i_n}))g(x_{i_1},\cdots,x_{i_n})) \\ &= f(x_1,\cdots,x_n)g(x_1,\cdots,x_n) \\ &= (fg)(x_1,\cdots,x_n).\end{aligned}$$

Therefore $fg \in K[x_1, \cdots, x_n]^G$. So, $K[x_1 \cdots x_n]^G \leq_R K[x_1 \cdots x_n]$.QED.

**Defintion:** A polynomial $f \in K[x_1 \cdots x_n]$ is **homogeneous of total degree** $k$ provided that every monomial appearing in $f$ has total degree $k$.

As an easy example of this definition, $f(x, y, z) = x^2y^2z + x^5 + y^2z^3$ is a homogeneous polynomial of degree 5. We follow with a theorem that connects the notion of homogenous and symmetric polynomials. A **homogeneous component** is the sum of all the monomials that share the same degree. For example $f(x, y, z)$ is a homogeneous component of

$g(x, y, z) = x^2y^2z + x^5 + y^2z^3 + xy + x^4yz + yz + xz.$

**Theorem:** A polynomial $f \in K[x_1 \cdots x_n]$ is symmetric if and only if all of its homogeneous components are symmetric.

Proof: Given a symmetric polynomial $f$, let $x_{i_1}, \cdots, x_{i_n}$ be a permutation of $x_1, \cdots, x_n$. This permutation takes a term of $f$ of total degree $k$ to one of the same total degree. Since $f(x_{i_1}, \cdots, x_{i_n}) = f(x_1, \cdots, x_n)$, it follows that the $k$-th homogenous component must also be symmetric. The converse is true since symmetric polynomials are closed under addition as seen in the proof above.QED.

The following theorem is important since we can, in some situations, only worry about homogeneous components which will simplify certain questions.

**Theorem:** A polynomial $f \in K[x_1 \cdots x_n]$ is invariant under a group $G \subset GL_n(K)$ if and only if its homogeneous components are.

Proof: $\Leftarrow$Let $f \in K[x_1 \cdots x_n]$ and the homogeneous components of $f$ be invariant under a group $G \subset GL_n(K)$. Then $f$ is invariant under $G$ since its homogeneous components are invariant, and we have already proved closure under addition.

$\Rightarrow$ As for the converse, let $f$ be invariant under $G$. Write $f = \sum_{1 \le k \le n} f_k$ where $f_k$ is the homogeneous component of degree $k$. The claim is $f_k(A\vec{x})$ is a homogeneous polynomial of degree $k$ for all $A \in G$. Therefore, it suffices to show if $x_1^{i_1} \cdots x_n^{i_n}$ is monomial of total degree $k = i_1 + \cdots + i_n$ then $(a_{1,1}x_1 + \cdots a_{1,n}x_n)^{i_1} \cdots (a_{n,1}x_1 + \cdots + a_{n,n}x_n)^{i_n}$ is a homogenous polynomial of degree $k$. The justification is that $f_k(A\vec{x})$ will be a homogeneous polynomial of degree $k$ since all the monomials change to homogeneous polynomials of degree $k$.

Let $A = (a_{i,j})$. Then,

$$\begin{aligned} (a_{1,1}x_1 + \cdots a_{1,n}x_n)^{i_1} \cdots (a_{n,1}x_1 + \cdots + a_{n,n}x_n)^{i_n} &= \left[\sum_{\alpha \in \mathbf{i_1}} b_{i_{1_\alpha}} X^\alpha\right] \cdots \left[\sum_{\alpha \in \mathbf{i_n}} b_{i_{n_\alpha}} X^\alpha\right] \\ &= \sum_{\alpha_j \in \mathbf{i_j}} (b_{i_{1_\alpha}} \cdots b_{i_{n_\alpha}}) X^{\alpha_1} \cdots X^{\alpha_n} \end{aligned}$$

where $\mathbf{i_j} = \{\alpha \in \mathbb{Z}^n_{\ge 0} : \alpha$ is the multinomial of a monomial in $(a_{j,1}x_1 + \cdots a_{j,n})^{i_n}$ expanded$\}$ and $b_{i_{j_\alpha}}$ is the respective coefficient. Finally, since $X^{\alpha_{i_j}}$ is of degree $j$, then we have $X^{\alpha_1} \cdots X^{\alpha_n}$ has degree $k$. The the theorem is proved.QED.

**Theorem:** Let $A_1, \cdots, A_m$ generate a group $G$. Then, $f \in K[x_1 \cdots x_n]^G$ if and only if $f(\vec{x}) = f(A_i\vec{x})$ for all $1 \le i \le m$.

Proof:$\Rightarrow$ Let $f \in K[x_1 \cdots x_n]^G$. Since $A_1, \cdots, A_m \in G$ then $f(\vec{x}) = f(A_i\vec{x})$ for all $1 \leq i \leq m$.

$\Leftarrow$ Let $f(\vec{x}) = f(A_i\vec{x})$ for all $1 \leq i \leq m$. Let $A \in G$, then $A = B_1 \cdots B_t$ where $B_i \in \{A_1, \cdots, A_m\}$ for each $i$. The proof precedes by induction. Then, $f(A\vec{x}) = f(B_1\vec{x}) = f(\vec{x})$ which is true by assumptions. Now suppose that $f(B_1 \cdots B_{t-1}\vec{x}) = f(\vec{x})$. Then,

$$\begin{aligned} f(A) &= f((B_1B_2 \cdots B_{t-1})B_t\vec{x}) \\ &= f(B_t\vec{x}) \\ &= f(\vec{x}). \end{aligned}$$

Therefore, $f \in K[x_1 \cdots x_n]^G$ and our theorem is proved.QED.

## 4.3 Examples, Solutions, and a New Theorem

We will first show that $K[x_1 \cdots x_n]^{S_n} = K[\mathbb{S}]$ by proving a well known theorem found in [2], and [4]. This theorem was initially proved by Gauss on his way to a second proof of the Fundamental Theorem of Algebra. The earliest known statement of lex order was for the theorem to follow.

**Fundamental Theorem of Symmetric Polynomials:** Every symmetric polynomial in $K[x_1 \cdots x_n]$ can be written uniquely as a polynomial in the elementary symmetric functions $\sigma_1, \cdots, \sigma_n$.

Proof: Let $f \in K[x_1 \cdots x_n]$ be any symmetric polynomial. Then the following algorithm rewrites $f$ uniquely as a polynomial in $\sigma_1, \cdots, \sigma_n$. We fix a monomial order $<$ to be graded lex order.

For any monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ occuring in the symmetric polynomial $f$ also all its images $x_{\sigma 1}^{\alpha_1} \cdots x_{\sigma_n}^{\alpha_n}$ under any permutation $\sigma$ of the variables occurring in $f$. This implies that $LT(f) = c \cdots x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_n^{\gamma_n}$ of $f$ satisfies $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$.

In our algorithm we now replace $f$ by the new symmetric polynomial
$\tilde{f} := f - c \cdot \sigma_1^{\gamma_1 - \gamma_2} \sigma_2^{\gamma_2 - \gamma_3} \cdots \sigma_{n-1}^{\gamma_{n-1} - \gamma_n} \sigma_n^{\gamma_n}$, we store the summand $c \cdot \sigma_1^{\gamma_1 - \gamma_2} \sigma_2^{\gamma_2 - \gamma_3} \cdots \sigma_{n-1}^{\gamma_{n-1} - \gamma_n} \sigma_n^{\gamma_n}$, and, if $\tilde{f}$ is nonzero, then we return to the beginning of the previous paragraph.

This process will terminate and here is why. By construction, the leading monomial of $c \cdot \sigma_1^{\gamma_1 - \gamma_2} \sigma_2^{\gamma_2 - \gamma_3} \cdots \sigma_{n-1}^{\gamma_{n-1} - \gamma_n} \sigma_n^{\gamma_n}$ equals $LT(f)$. Hence, in the difference defining $\tilde{f}$ the two leading monomial cancel, and we get $LT(\tilde{f}) < LT(f)$. The set of monomials $\{m : m < LT(f)\}$ is finite because their degree is bounded and by well ordering. Thus, the algorithm must terminate.

For uniqueness, we will fix the monomial order to be lex. So suppose, in $K[x_1 \cdots x_n]$ we have $g_1(\sigma_1, \cdots, \sigma_n) = g_2(\sigma_1, \cdots, \sigma_n)$ and then define $g = g_1 - g_2$. For a contradiction, suppose $g \neq 0$ in $K[y_1 \cdots y_n]$. If we write $g = \sum_\beta a_\beta y^\beta$, then $g(\sigma_1, \cdots, \sigma_n)$ is a sum of the polynomials $g_\beta = a_\beta \sigma_1^{\beta_1} \cdots \sigma_n^{\beta_n}$, where $\beta = (\beta_1, \cdots, \beta_n)$. Furthermore, $LT(g_\beta) = a_\beta x_1^{\beta_1 + \cdots \beta_n} x_2^{\beta_2 + \cdots \beta_n} \cdots x_n^{\beta_n}$. Also note that $(\beta_1, \cdots, \beta_n) \mapsto (\beta_1 + \cdots + \beta_n, \beta_2 + \beta_n, \cdots, \beta_n)$ is injective. Thus, the $g_\beta$'s have distinct leading terms. If we pick $\beta$ such that $LT(g_\beta) > LT(g_\gamma)$ for all $\gamma \neq \beta$, then $LT(g_\beta)$ will be greater than all monomial terms of the $g_y$'s. It follows that nothing can cancel with $LT(g_\beta)$ and we arrive at the contradiction $g(\sigma_1, \cdots, \sigma_n) \neq 0$. QED.

With this we can conclude $K[x_1 \cdots x_n]^{S_n} = K[\mathbb{S}]$ and each $f \in K[x_1 \cdots x_n]^{S_n}$ is written in terms of $\mathbb{S}$ uniquely. We will now follow with computer implementation using *Mathematica.* The command *SymmetricReduction* will rewrite polynomials in terms of sig$k$ where sig$k := \sigma_k$.

$In[1] := f = x1 * x2 * x3 * x4^2 + x1 * x2 * x3^2 * x4 + x1 * x2^2 * x3 * x4 + x1^2 * x2 * x3 * x4;$

In[2]:= SymmetricReduction[f, {x1, x2, x3, x4}, {sig1, sig2, sig3, sig4}]

Out[2]= {sig1 sig4, 0}.

The output is an ordered pair and 0 to the right of the comma in the output signifies $f$ is a symmetric polynomial. Moreover, the output also tells us $f = \sigma_1 \cdot \sigma_4$.

To show that generators of $K[x_1 \cdots x_n]^{S_n}$ are not unique we will prove $\mathbb{P}$ is also a viable set of generators. This result is found in [2].

**Theorem:** Let $\mathbb{Q} \subset K$. Every symmetric polynomial in $k[x_1 \cdots x_n]$ can be written as polynomials in the power sums $S_k$ for $1 \leq k \leq n$.

Proof: From the Fundamental Theorem of Symmetric Polynomials it suffices to show that every element of $\mathbb{S}$ can be written in terms of elements of $\mathbb{P}$. We will now introduce the well-known Newton Identities namely

$$\begin{array}{rcl} s_k - \sigma_1 s_{k-1} + \cdots + (-1)^k \sigma_{k-1} s_1 + (-1)^k k \sigma_k & = & 0, 1 \leq k \leq n \\ s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{n-1} \sigma_{n-1} s_{k-n+1} + (-1) \sigma_n s_{k-n} & = & 0, k > 0. \end{array}$$

To complete this proof. We will now proceed by induction. For $k = 1$, $s_1$ and $\sigma_1$ are defined to be equal summands.

For the inductive hypothesis assume our claim is proved for $1, 2, \cdots, k-1$, then directly from the Newton identities we see that

$$\sigma_k = (-1)^{k-1}\frac{1}{k}(s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1).$$

We can divide by $k$ since $\mathbb{Q} \subset K$. Now, by our inductive hypothesis, we have prove that $\sigma_k$ can be written in terms of elements in $\mathbb{P}$, thus our theorem is proved.QED.

As an example, we find generators for $K[xy]^{K_4}$, where $K_4 = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle$. ($K_4$ is actually the Klein-4 group.) This example can be found on page 332 of [2].

First note that $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. So $f \in K[xy]^{K_4}$ if and only if $f(x,y) = f(-x,y) = f(x,-y)$. Now we will write $f = \sum_{ij} a_{ij}x^iy^j$. Then we have the following

$$\begin{aligned} f(x,y) &= f(-x,y) \\ \Leftrightarrow \sum_{ij} a_{i,j}x^iy^j &= \sum_{ij} a_{i,j}(-x)^iy^j \\ \Leftrightarrow \sum_{ij} a_{i,j}x^iy^j &= \sum_{ij} a_{i,j}(-1)^ix^iy^j \\ \Leftrightarrow a_{i,j} &= (-1)^ia_{i,j} \text{ for all } i,j \\ \Leftrightarrow a_{i,j} &= 0 \text{ when } i \text{ is odd.} \end{aligned}$$

Similarly,

$$\begin{aligned} f(x,y) &= f(x,-y) \\ \Leftrightarrow \sum_{ij} a_{i,j}x^iy^j &= \sum_{ij} a_{i,j}x^i(-y)^j \\ \Leftrightarrow \sum_{ij} a_{i,j}x^iy^j &= \sum_{ij} a_{i,j}(-1)^jx^iy^j \\ \Leftrightarrow a_{i,j} &= (-1)^ja_{i,j} \text{ for all } i,j \\ \Leftrightarrow a_{i,j} &= 0 \text{ when } j \text{ is odd.} \end{aligned}$$

Thus $f \in K[xy]^{K_4}$ if and only if $f$ can be written in terms of $x^2$ and $y^2$. Therefore, we may conclude $K[xy]^{K_4} = K[\mathbb{F}]$ where $\mathbb{F} := \{x^2, y^2\}$.

For another example we will find a set of generators for $K[xy]^{G_1}$ where $G_1 = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \right\rangle$. So now, doing a similar procedure, $f \in K[xy]$ if and only if $f(x,y) = f(2x,2y)$. Then we will write $f = \sum_{ij} a_{i,j}x^iy^j$. So,

$$\begin{array}{rcl} f(x,y) & = & f(2x,2y) \\ \Leftrightarrow \displaystyle\sum_{ij} a_{i,j}x^iy^j & = & \displaystyle\sum_{ij} a_{i,j}(2x)^i(2y)^j \\ \Leftrightarrow \displaystyle\sum_{ij} a_{i,j}x^iy^j & = & \displaystyle\sum_{ij} 2^{i+j}a_{i,j}x^iy^j \\ \Leftrightarrow a_{i,j} & = & 2^{i+j}a_{i,j} \\ \Leftrightarrow a_{i,j} = 0 & \text{or} & 2^{i+j} = 1 \end{array}$$

So $i + j = 0$ which implies $i = j = 0$ since $i, j \geq 0$. Thus $f \in K[xy]^{G_1}$ if and only if $f$ is constant. That is, $K[xy]^{G_1} = K[\mathbb{F}_1]$ where $\mathbb{F}_1 = \{1\}$.

Now for another less trivial example, we will find a set of generators for $K[xy]^{G_2}$ where $G_2 = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \right\rangle$.So now $f \in K[xy]$ if and only if $f(x,y) = f(2x, \frac{1}{2}y)$. Then we will write $f = \sum_{ij} a_{i,j}x^iy^j$. So,

$$\begin{array}{rcl} f(x,y) & = & f(2x,\frac{1}{2}y) \\ \Leftrightarrow \displaystyle\sum_{ij} a_{i,j}x^iy^j & = & \displaystyle\sum_{ij} a_{i,j}(2x)^i(\frac{1}{2}y)^j \\ \Leftrightarrow \displaystyle\sum_{ij} a_{i,j}x^iy^j & = & \displaystyle\sum_{ij} 2^{i-j}a_{i,j}x^iy^j \\ \Leftrightarrow a_{i,j} & = & 2^{i-j}a_{i,j} \\ \Leftrightarrow a_{i,j} = 0 & \text{or} & 2^{i-j} = 1 \end{array}$$

So $i - j = 0$ which implies $i = j$ since $i, j$. Thus $f \in K[xy]^{G_1}$ if and only if $f$ is written in terms of $xy$. That is, $K[xy]^{G_1} = K[\mathbb{F}_2]$ where $\mathbb{F}_2 = \{xy\}$.

Let's take a closer look at $K_4, G_1$, and $G_2$. First, $K_4 = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$ is finite.

For $G_1 = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^j : j \in \mathbb{Z} \right\} = \left\{ \begin{pmatrix} 2^j & 0 \\ 0 & 2^j \end{pmatrix} : j \in \mathbb{Z} \right\}$. It is true $\begin{pmatrix} 2^{j_1} & 0 \\ 0 & 2^{j_1} \end{pmatrix} = \begin{pmatrix} 2^{j_2} & 0 \\ 0 & 2^{j_2} \end{pmatrix}$ implies $j_1 = j_2$ and $\begin{pmatrix} 2^{j_1} & 0 \\ 0 & 2^{j_1} \end{pmatrix}\begin{pmatrix} 2^{j_2} & 0 \\ 0 & 2^{j_2} \end{pmatrix} = \begin{pmatrix} 2^{j_1+j_2} & 0 \\ 0 & 2^{j_1+j_2} \end{pmatrix}$

Likewise for $G_2 = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}^j : j \in \mathbb{Z} \right\} = \left\{ \begin{pmatrix} 2^j & 0 \\ 0 & \frac{1}{2}^j \end{pmatrix} : j \in \mathbb{Z} \right\}$. It is true $\begin{pmatrix} 2^{j_1} & 0 \\ 0 & \frac{1}{2}^{j_1} \end{pmatrix} = \begin{pmatrix} 2^{j_2} & 0 \\ 0 & \frac{1}{2}^{j_2} \end{pmatrix}$ implies $j_1 = j_2$ and $\begin{pmatrix} 2^{j_1} & 0 \\ 0 & \frac{1}{2}^{j_1} \end{pmatrix} \begin{pmatrix} 2^{j_2} & 0 \\ 0 & \frac{1}{2}^{j_2} \end{pmatrix} = \begin{pmatrix} 2^{j_1+j_2} & 0 \\ 0 & \frac{1}{2}^{j_1+j_2} \end{pmatrix}$.

Thus $\phi_1 : \mathbb{Z} \mapsto G_1$ defined by $\phi_1(n) = \begin{pmatrix} 2^n & 0 \\ 0 & 2^n \end{pmatrix}$ and $\phi_2 : \mathbb{Z} \mapsto G_2$ defined by $\phi_2(n) = \begin{pmatrix} 2^n & 0 \\ 0 & \frac{1}{2}^n \end{pmatrix}$ are both isomorphisms. So $G_1 \cong \mathbb{Z} \cong G_2$. The first point is that $G_1$ and $G_2$ are infinite matrix groups. The second point is that even though these group are isomorphic, the associated generators were different, that is $\mathbb{F}_1 \neq \mathbb{F}_2$. The following is a new theorem that was motivated by the examples above.

**Theorem:** Let $n \in \mathbb{Z}^+ \cup \{0\}$. For all $m \geq n$ there exists $G \in GL_m(K)$ where $G \cong \mathbb{Z}$ and $K[x_1, \cdots, x_m]^G \cong K[x_1, \cdots, x_n]$.

Proof: Let $n \in \mathbb{Z}^+ \cup \{0\}$ and $m \geq n$. Choose primes with $p_1 < p_2 < \cdots < p_{m-n}$. Let

$$G = \left\langle \begin{pmatrix} \frac{1}{p_1} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{p_2} & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{1}{p_{m-n}} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \Pi_{i=1}^{m-n} p_i & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \right\rangle.$$

Now the claim is $K[x_1, \cdots, x_m]^G = K[x_1, \cdots, x_n]$. To see that note that $f \in K[x_1, \cdots, x_m]^G$ if and only if $f(\frac{1}{p_1}x_1, \frac{1}{p_2}x_2, \cdots, \frac{1}{p_{m-n}}x_{m-n}, \Pi_{i=1}^{m-n} p_i x_{m-n+1}, x_{m-n+2}, \cdots, x_m) = f(x_1, \cdots, x_m)$. Then

$$\begin{aligned} & \sum_{i_1, \cdots, i_m} a_{i_1, \cdots, i_m} (\frac{1}{p_1}x_1)^{i_1} (\frac{1}{p_2}x_2)^{i_2} \cdots (\frac{1}{p_{m-n}}x_{m-n})^{i_{m-n}} (\Pi_{i=1}^{m-n} p_i x_{m-n+1})^{i_{m-n+1}} (x_{m-n+2})^{i_{m-n+2}} \cdots (x_m)^{i_m} \\ = & \sum_{i_1, \cdots, i_m} a_{i_1, \cdots, i_m} (x_1)^{i_1} (x_2)^{i_2} \cdots (x_{m-n})^{i_{m-n}} (x_{m-n+1})^{i_{m-n+1}} (x_{m-n+2})^{i_{m-n+2}} \cdots (x_m)^{i_m} \\ \Leftrightarrow & \; a_{i_1, \cdots, i_m} (\frac{1}{p_1})^{i_1} (\frac{1}{p_2})^{i_2} \cdots (\frac{1}{p_{m-n}})^{i_{m-n}} (\Pi_{i=1}^{m-n} p_i)^{i_{m-n+1}} = a_{i_1, \cdots, i_m} \\ \Leftrightarrow & \; a_{i_1, \cdots, i_m} = 0 \text{ or } i_1 = \cdots = i_{m-n+1}. \end{aligned}$$

So $K[x_1, \cdots, x_m]^G = K[x_1 \cdots x_{m-n+1}, x_{m-n+2}, \cdots, x_m]$. To check to make sure we have the correct number of indeterminates, note that $m-(m-n+1)+1 = n$. Thus $K[x_1, \cdots, x_m]^G \cong K[x_1, \cdots, x_n]$. Finally, since $G$ is generated by a diagonal matrix, $G \cong \mathbb{Z}$. Q.E.D.

# Chapter 5

# Groebner Bases and Invariant Theory

## 5.1 Invariant Rings and Decidability Theorems

These results and proofs can be found in [2]. The following theorem gives a way to rewrite a symmetric polynomial in $K[x_1, \cdots, x_n]$ in terms of the elementary symmetric polynomials. As we will see, an analogous result exists for $GL_n(K)$.

**Theorem:** In the ring $K[x_1 \cdots x_n y_1 \cdots, y_n]$, fix a monomial order where any monomial involving one of $x_1, \cdots, x_n$ is greater than all monomials in $K[y_1 \cdots y_n]$. Let $G$ be a Groebner basis of the ideal $\langle \sigma_1 - y_1, \cdots, \sigma_n - y_n \rangle \subset K[x_1 \cdots x_n y_1 \cdots y_n]$. Given $f \in K[x_1 \cdots x_n]$. Then:

(i) $f$ is symmetric if and only if $\bar{f}^G \in K[y_1 \cdots y_n]$.

(ii) If $f$ is symmetric, then $f = \bar{f}^G(\sigma_1, \cdots, \sigma_n)$ is the unique expression of $f$ as a polynomial in the elementary symmetric polynomials in $\sigma_1, \cdots, \sigma_n$.

Proof: In the ring $K[x_1, \cdots, x_n, y_1, \cdots, y_n]$, fix a monomial order where any monomial involving one of $x_1, \cdots, x_n$ is greater than all monomials in $K[y_1, \cdots, y_n]$. Let $G = \{g_1, \cdots, g_t\}$ be a Groebner basis of the ideal $\langle \sigma_1 - y_1, \cdots, \sigma_n - y_n \rangle \subset K[x_1 \cdots x_n y_1 \cdots y_n]$. Let $f \in K[x_1 \cdots x_n]$. Then, after division by $G$,

$$f \;\; = \;\; h_1 g_1 + \cdots + h_t g_t + \bar{f}^G \text{ where } h_1, \cdots, h_t \in K[x_1 \cdots x_n y_1 \cdots y_n].$$

We may assume $g \neq 0$ for all $g \in G$.

First for (i). First suppose $\bar{f}^G \in K[y_1 \cdots y_n]$. Let $y_i := \sigma_i$ for each $i$ in the formula above. Note that $f$ will not change since $f$ is a function of the indeterminates $x_1, \cdots, x_n$. Now, $\langle \sigma_1 - y_1, \cdots, \sigma_n - y_n \rangle = \langle 0 \rangle$, thus $g_1 = \cdots = g_t = 0$. So we can now see that

$f = \bar{f}^G(\sigma_1, \cdots, \sigma_n)$. In other words, $f$ is symmetric.

Let $f \in K[x_1 \cdots x_n]$ be symmetric. Then $f = g(\sigma_1, \cdots, \sigma_n)$ for some $g \in K[y_1 \cdots y_n]$. We want to show $g = \bar{f}^G$. First we have in $K[x_1 \cdots x_n y_1 \cdots y_n]$

$$\begin{aligned} \sigma_1^{\alpha_1} \cdots \sigma_n^{\alpha_n} &= (y_1 + (\sigma_1 - y_1)^{\alpha_1} \cdots (y_n + (\sigma_n - y_n)^{\alpha_n} \\ &= y_1^{\alpha_1} \cdots y_n^{\alpha_n} + B_1(\sigma_1 - y_1) + \cdots + B_n(\sigma_n - y_n) \end{aligned}$$

for some $B_1, \cdots, B_n \in K[x_1 \cdots x_n y_1 \cdots y_n]$. Then $g(\sigma_1, \cdots, \sigma_n)$ can be written in the monomials given above. Thus

$$f = g(\sigma_1, \cdots, \sigma_n) = C_1(\sigma_1 - y_1) + \cdots + C_n(\sigma_n - y_n) + g(y_1, \cdots y_n)$$

where $C_1, \cdots C_n \in K[x_1, \cdots, x_n, y_1, \cdots, y_n]$ by grouping in an appropriate way.

For a contradiction, suppose there is a term of $g$ is divisible by an element of $LT(G)$, that is suppose for some $i$, $LT(g_i)$ divides a term of $g$. This immediately implies $g_i \in K[y_1 \cdots y_n]$ by our choice of monomial order and since $g \in K[y_1, \cdots, y_n]$. Now define $y_i := \sigma_i$. Since $g_i \in \langle \sigma_1 - y_1, \cdots, \sigma_n - y_n \rangle$, we have already seen that $g_i$ equals zero after f the substitution given above. Then $g_i \in K[y_1 \cdots y_n]$ means $g_i(\sigma_1, \cdots, \sigma_n) = 0$. Thus, the uniqueness guaranteed by the Fundamental Theorem of Symmetric Polynomials implies that $g_i = 0$, which is a contradiction. So, no term of $g$ is divisible by an element of $LT(G)$. Thus, by the division algorithm, $g = \bar{f}^G$.

As for (ii), this is true by the construction of (i).QED.

The following theorem is analogous to the one above and shows that we can use Groebner bases to rewrite polynomials in terms of the generators.
**Theorem:** Suppose that $f_1, \cdots f_m \in K[x_1 \cdots x_n]$ are given. Fix a monomial order where any monomial involving one of $x_1, \cdots, x_n$ is greater than all monomials in $K[y_1 \cdots y_n]$. Let $G$ be a Groebner basis of the ideal $\langle f_1 - y_1, \cdots, f_n - y_n \rangle \subset K[x_1 \cdots x_n y_1 \cdots y_n]$. Let $f \in K[x_1 \cdots x_n]$. Then:

(i) $f \in K[f_1 \cdots f_m]$ if and only if $\bar{f}^G \in K[y_1 \cdots y_n]$.

(ii) If $f$ is symmetric, then $f = \bar{f}^G(f_1, \cdots, f_n)$ is the unique expression of $f$ as a polynomial in the elementary symmetric polynomials in $f_1, \cdots, f_n$.

Proof: In the ring $K[x_1, \cdots, x_n, y_1, \cdots, y_m]$, fix a monomial order where any monomial involving one of $x_1, \cdots, x_n$ is greater than all monomials in $K[y_1, \cdots, y_m]$. Let $G = \{g_1, \cdots, g_t\}$ be a Groebner basis of the ideal $\langle f_1 - y_1, \cdots, f_m - y_m \rangle \subset K[x_1, \cdots, x_n, y_1, \cdots, y_m]$. Let $f \in K[x_1 \cdots x_n]$. Then, after division by $G$,

$$f = h_1 g_1 + \cdots + h_t g_t + \bar{f}^G \text{ where } h_1, \cdots, h_t \in K[x_1, \cdots, x_n, y_1, \cdots, y_m].$$

We may assume $g \neq 0$ for all $g \in G$.

First for (i). First suppose $\bar{f}^G \in K[y_1 \cdots y_m]$. Let $y_i := f_i$ for each $i$ in the formula above. Now $f$ will not change since $f$ is the indeterminates $x_1, \cdots, x_n$. Now, $\langle f_1 - y_1, \cdots, f_m - y_m \rangle = \langle 0 \rangle$, thus $g_1 = \cdots = g_t = 0$. So we can now see that $f = \bar{f}^G(f_1, \cdots, f_m)$, in other words $f \in K[f_1 \cdots f_m]$.

Let $f \in K[f_1 \cdots f_m]$. Then $f = g(f_1, \cdots, f_n)$ for some $g \in K[y_1 \cdots y_n]$. We want to show $g = \bar{f}^G$. First we have in

$$\begin{aligned}\sigma_1^{\alpha_1} \cdots \sigma_n^{\alpha_n} &= (y_1 + (\sigma_1 - y_1)^{\alpha_1} \cdots (y_n + (\sigma_n - y_n)^{\alpha_n} \\ &= y_1^{\alpha_1} \cdots y_n^{\alpha_n} + B_1(\sigma_1 - y_1) + \cdots + B_n(\sigma_n - y_n)\end{aligned}$$

for some $B_1, \cdots, B_n \in K[x_1, \cdots, x_n, y_1, \cdots, y_n]$. Then $g(\sigma_1, \cdots, \sigma_n)$ can be written in the monomials given above. Thus

$$f = g(\sigma_1, \cdots, \sigma_n) = C_1(\sigma_1 - y_1) + \cdots + C_n(\sigma_n - y_n) + g(y_1, \cdots y_n)$$

where $C_1, \cdots C_n \in K[x_1 \cdots x_n y_1 \cdots y_n]$ by grouping in an appropriate way.

Unlike before, $\bar{f}^G$ need not equal $g$. To remedy this, let $G' = G \cap K[y_1 \cdots y_m]$. Thus $G' = \{g'_1, \cdots, g'_s\}$. Dividing $g$ by $G'$ we find

$$g = D_1 g_1 + \cdots + D_s g_s + g' \text{ where } B_1, \cdots, B_s, g' \in K[y_1 \cdots y_m].$$

Since for each $i$ $g_i \in \langle f_1 - y_1, \cdots f_m - y_m \rangle$, we can combine the two previous equation and fine

$$f = C'_1(f_1 - y_1) + \cdots + C'_m(f_m - y_m) + g'(y_1, \cdots y_m).$$

The claim is $g' = \bar{f}^G$. For a contradiction and using the division algorithm, suppose there is a term of $g'$ is divisible by an element of $LT(G)$, that is suppose, for some $i$, $LT(g_i)$ divides a term of $g'$. This immediately implies $g_i \in K[y_1 \cdots y_m]$ by our choice of monomial order and $g' \in K[y_1, \cdots, y_m]$. So $g_i \in G'$. Since $g'$ is a remainder on division by $G'$, $LT(g_i)$ cannot divide any term of $g'$ which is a contradiction. So, $g' = \bar{f}^G$.

As for (ii), this is true by the construction of (i).QED.

We have two theorems which are implemented in *Mathematica* in the next section. The results in the section are both intriguing and computationally useful.

## 5.2 Decidability Algorithms Based on the Previous Theorems

The first Algorithm is for the first Theorem in the previous section. We define a function *gbsym* and $f$ is the polynomial we wont to rewrite and $n$ is the number of variables.

```
gbsym[f_, n_]:=Do[{X = Union[Array[a, n], Array[sym, n]],

F = {},

Do[{F = Union[{SymmetricPolynomial[j, Array[a, n]] − sig[j]}, F]}, {j, 1, n}],

{b, c} = PolynomialReduce[f, GroebnerBasis[F, X], X];

Print[c]}, {k, 0, 0}];
```

```
gbsym[a[1]^10 + a[2]^10 + a[3]^10, 3]
```

$\text{sig}[1]^{10}-10\text{sig}[1]^{8}\text{sig}[2]+35\text{sig}[1]^{6}\text{sig}[2]^{2}-50\text{sig}[1]^{4}\text{sig}[2]^{3}+25\text{sig}[1]^{2}\text{sig}[2]^{4}-2\text{sig}[2]^{5}+10\text{sig}[1]^{7}\text{sig}[3]-60\text{sig}[1]^{5}\text{sig}[2]\text{sig}[3]+100\text{sig}[1]^{3}\text{sig}[2]^{2}\text{sig}[3]-40\text{sig}[1]\text{sig}[2]^{3}\text{sig}[3]+25\text{sig}[1]^{4}\text{sig}[3]^{2}-60\text{sig}[1]^{2}\text{sig}[2]\text{sig}[3]^{2}+15\text{sig}[2]^{2}\text{sig}[3]^{2}+10\text{sig}[1]\text{sig}[3]^{3}$

The output above rewrites $a[1]^{10} + a[2]^{10} + a[3]^{10}$ in terms of the elementary symmetric polynomials. The next Algorithm is for the second theorem in the previous section. We define a function *gbinv*. Where $f$ is the polynomial we want to write in terms of $F$ in $n$ variables.

```
gbinv[f_, n_, F_]:=

Do[{X = Union[Array[a, n], Array[y, Length[F]]], S = {},

Do[S = Union[S, {inv[j]}], {j, 1, Length[F]}],

{b, c} = PolynomialReduce[f, GroebnerBasis[F − S, X], X];

Print[c]}, {k, 0, 0}]
```

```
F = {a[1]^2 + a[2]^2, a[1]^3a[2] − a[1]a[2]^3, a[1]^2a[2]^2};
```

```
gbinv[a[1]^8 + 2 * a[1]^6 * a[2]^2 - a[1]^5 * a[2]^3 + 2 * a[1]^4a[2]^4 + a[1]^3a[2]^5+
2 * a[1]^2a[2]^6 + a[2]^8, 2, F]
```

$\text{inv}[1]^4 - 2\text{inv}[1]^2\text{inv}[3] - \text{inv}[2]\text{inv}[3]$

```
gbinv[a[1]^2 + a[2]^2, 2, F]
```

$\text{inv}[1]$

We give two samples of output above to verify the algorithm works. So now we can computationally, using Groebner bases, rewrite polynomials in terms of the invariant generators for $S_n$ and finite subgroups of $GL_n(K)$.

## 5.3 Finding Invariants

The next definition and theorem can be found in [7].

**Definition:** Given a finite matrix group $G \subset GL_n(K)$, the **Reynolds operator** of $G$ is the map $R_G : K[x_1 \cdots x_n] \to K[x_1 \cdots x_n]$ defined by $REY_G(f)(\vec{x}) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \vec{x})$ for $f(\vec{x}) \in K[\vec{x}]$.

**Theorem:** The Reynolds operator on a finite group has the following properties:

i)$R_G(\lambda f + vg) = \lambda R_G(f) + vR_G(g)$ for all $f, g \in K[x_1 \cdots x_n]$ and $\lambda, v \in K$;
ii)$R_G|_G$ is the identity map;
iii) $R_G(fI) = R_G(f) \cdot I$ for all $f \in K[x_1 \cdots x_n]$ and $I \in K[x_1 \cdots x_n]^G$.

Proof: Let $G \subset GL_n(K)$ be a finite group.

Let $f, g \in K[x_1 \cdots x_n]$ and $\lambda, v \in K$. Then

$$\begin{aligned} REY_G(\lambda f + vg)(\vec{x}) &= \frac{1}{|G|} \sum_{A \in G} (\lambda f + vg)(A \cdot \vec{x}) \\ &= \frac{1}{|G|} \sum_{A \in G} \lambda f(A \cdot \vec{x}) + vg(A \cdot \vec{x}) \\ &= \frac{\lambda}{|G|} \sum_{A \in G} f(A \cdot \vec{x}) + \frac{v}{|G|} \sum_{A \in G} f(A \cdot \vec{x}) \\ &= \lambda REY_G(f) + vREY_G(g). \end{aligned}$$

So $i)$ is proven true.

Let $f \in K[\vec{x}]^G$ then $f(A \cdot \vec{x}) = f(\vec{x})$ for all $A \in G$. Then we see that

$$REY_G(f)(\vec{x}) = \frac{1}{|G|}\sum_{A\in G} f(A\cdot\vec{x}) = \frac{1}{|G|}\sum_{A\in G} f(\vec{x}) = f(\vec{x}).$$

So $ii)$ is proved.

Let $f \in K[x_1 \cdots x_n]$ and $I \in K[x_1 \cdots x_n]^G$. Then

$$\begin{aligned} REY_G(fI)(\vec{x}) &= \frac{1}{|G|}\sum_{A\in G}(fI)(A\cdot\vec{x}) \\ &= \frac{1}{|G|}\sum_{A\in G}(f(A\cdot\vec{x}))(I(A\cdot\vec{x})) \\ &= \frac{1}{|G|}\sum_{A\in G}(f(A\cdot\vec{x}))(I(\vec{x})) \\ &= I(\vec{x})\frac{1}{|G|}\sum_{A\in G}(f(A\cdot\vec{x})) \\ &= (I\cdot REY_G(f))(\vec{x}). \end{aligned}$$

So, iii) is proved and we now know that $G$ is reductive. QED.

A nonconstructive theorem that proves finite generation for $K[\vec{x}]^G$ when $G$ is finite will be presented which uses the Hilbert Basis Lemma. This result can be found in [4] and [7].

**Hilbert Finiteness Theorem:** The invariant ring $K[\vec{x}]^G$ of a finite matrix group $G \subset GL_n(K)$ is finitely generated.

Proof: Let $K[\vec{x}]^G_+$ denote the set of all homogeneous invariants of positive degree. From the Hilbert Basis (Theorem) Lemma we see that $\langle K[\vec{x}]^G_+\rangle = \langle I_1, \cdots, I_m\rangle$ for some $I_1, \cdots, I_m \in K[\vec{x}]^G_+$. Now it will be shown that $K[\vec{x}]^G = K[I_1, \cdots, I_m]$.

For a contradiction suppose $I \in K[\vec{x}]^G - K[I_1, \cdots, I_m]$ and where $I$ is homogeneous of minimum degree. Then $I = \sum_{i=1}^m f_i I_i$ since $I \in \langle I_1, \cdots, I_m\rangle$ for homogeneous polynomials $f_i \in K[\vec{x}]$ and $\deg(f_i) < \deg(I)$. Now we see applying the Reynolds operator

$$I = REY_G(I) = REY_G\left(\sum_{i=1}^m f_i I_i\right) = \sum_{i=1}^m REY_G(f_i) I_i.$$

Note that $\deg(REY_G(f_i)) = \deg(f_i) < \deg(I)$. This means $REY_G(f_i) \in K[I_1, \cdots, I_m]$ and as a consequence $I \in K[I_1, \cdots, I_m]$ which is a contradiction. So $K[\vec{x}]^G \subset K[I_1, \cdots, I_m]$ since we proved it suffices to only consider homogeneous polynomials. Also, $K[I_1, \cdots, I_m] \subset K[\vec{x}]^G$ by construction. QED.

Since any $I \in K[\vec{x}]^G$ takes on the form $I = \sum_{i=1}^m REY_G(f_i)I_i$, as we see in the proof, and by the linearity of $REY_G$ $I$ can be written in terms of $\{REY_G(x^\alpha) : \alpha \in \mathbb{Z}^n_{\geq 0}\}$. A theorem proved by Noether provides a way to find all the invariants algorithmically. This new process does not yet use Groebner bases and is inefficient for large groups. Moreover, redundancies are likely to exist. This result is found in [2] and [7].

**Noether's Degree Bound:** Given a finite matrix group $G \subset GL_n(K)$, we have $k[\vec{x}]^G = K[REY_G(x^\beta) : |\beta| \leq |G|]$.

Proof: First note that $(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} a_\alpha x^\alpha$ where $a_\alpha \in \mathbb{Z}^+$.

Let $A_i$ denote the $i$th row of $A$ so $A_i\vec{x} = \sum_{j=1}^n a_{i,j}x_j$. Let $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ and define $(A\vec{x})^\alpha := \Pi_{i=1}^n (A_i\vec{x})^{\alpha_i}$. Now we can see

$$\begin{aligned} REY_G(x^\alpha) &= \frac{1}{|G|}\sum_{A\in G}\left[\Pi_{i=1}^n(a^A_{i,1}x_1 + \cdots + a^A_{i,n}x_n)^{\alpha_i}\right] \\ &= \frac{1}{|G|}\sum_{A\in G}[\Pi_{i=1}^n (A_i\vec{x})^{\alpha_i}] \\ &= \frac{1}{|G|}\sum_{A\in G}(A\vec{x})^\alpha. \end{aligned}$$

Now we will introduce the new indeterminates $u_1, \cdots, u_n$ and substitute $u_iA_i\vec{x}$ for each $x_i$. Then we have

$$\begin{aligned} (u_1A_1x_1 + \cdots + u_nA_nx_n)^k &= \sum_{|\alpha|=k} a_\alpha(u_1A_1\vec{x})^{\alpha_1}\cdots(u_nA_n\vec{x})^{\alpha_n} \\ &= \sum_{|\alpha|=k} a_\alpha(A\vec{x})^\alpha u^\alpha. \end{aligned}$$

Then

$$\begin{aligned} S_k : &= \sum_{A\in G}(u_1A_1\vec{x} + \cdots + u_1A_n\vec{x})^k \\ &= \sum_{|\alpha|=k} a_\alpha\left(\sum_{A\in G}(A\vec{x})^\alpha\right)u^\alpha \\ &= \sum_{|\alpha|=k}|G|a_\alpha REY_G(x^\alpha)u^\alpha. \end{aligned}$$

Now let $U_A = u_1A_1\vec{x} + \cdots + u_1A_n\vec{x}$ where $A_i$ is the $i$th row of $A \in G$. Any symmetric function in terms of the $U_A$ can be written in terms of $S_k$ by a previous theorem. Now $S_k$ is a symmetric polynomial in terms of $U_A$, so $S_k = F(S_1, \cdots, S_{|G|})$. Now substituting we find

$$\sum_{|\alpha|=k} |G| a_\alpha REY_G(x^\alpha) u^\alpha = F\left( \sum_{|\beta|=1} |G| a_\alpha REY_G(x^\beta) u^\beta, \cdots, \sum_{|\beta|=|G|} |G| a_\alpha REY_G(x^\beta) u^\beta \right)$$

Expanding the right side and equating the coefficients of $u^\alpha$, it follows that $|G|a_\alpha REY_G(x^\alpha)$ is a polynomial in $REY_G(x^\beta)$ where $|\beta| \leq |G|$. Since $k$ has characteristic zero, then $|G|a_\alpha$ is nonzero, and hence $REY_G(x^\alpha)$ has the desired form.QED

The implementation presented above is not efficient if $G$ is large, but it will work and terminate based on the theorems above.

## 5.4 Finding Invariants with Groebner Bases

This entire section is pulled from [7]. We are looking for a set of fundamental invariants $\{\theta_1, \cdots, \theta_n, \eta_1, \cdots, \eta_t\}$ where each $I \in \mathbb{C}[\vec{x}]$ can be written uniquely as $I(\vec{x}) = \sum_{i=1}^t \eta_i(\vec{x}) \cdot p_i(\theta_1(\vec{x}), \cdots, \theta_n(\vec{x}))$ where $p_1, \cdots, p_t$ are $n$-variant polynomials. We will define the Hironaka decomposition $\bigoplus_{i=1}^t \eta_i \cdot \mathbb{C}[\theta_1, \cdots, \theta_n]$ to be the set of polynomials of the form $\sum_{i=1}^t \eta_i(\vec{x}) \cdot p_i(\theta_1(\vec{x}), \cdots, \theta_n(\vec{x}))$ where $p_1, \cdots, p_t$ are $n$-variant polynomials. We will call $\theta_1, \cdots, \theta_n$ to be primary invariants and $\eta_1, \cdots, \eta_t$ be secondary invariants. So to restate our goal given a finite group $G$ we want to find $\{\theta_1, \cdots, \theta_n, \eta_1, \cdots, \eta_t\}$ so $\mathbb{C}[\vec{x}]^G = \bigoplus_{i=1}^t \eta_i \cdot \mathbb{C}[\theta_1, \cdots, \theta_n]$.

Let's take $G = \{1\}$ then $\mathbb{C}[x]^G = \mathbb{C}[x] = \mathbb{C}[x^2] \bigoplus x\mathbb{C}[x^2]$. The first equality is a result from previous sections and the second is a direct result from the paragraph above. This establishes the fact that Hironaka decompositions are not unique for $G$ in general and in fact the example given before has infinitely many decompositions.

Now that us prove a lemma that will be used in an algorithm for finding $\{\theta_1, \cdots, \theta_n, \eta_1, \cdots, \eta_t\}$.

**Lemma:** Let $G \subset GL_n(\mathbb{C})$ be any finite matrix group, and let

$$I^G = \langle I \in \mathbb{C}[\vec{x}] : I(A \cdot \vec{x}) = I(\vec{x}) \text{ for all } A \in G, \deg I > 0 \rangle$$

Then $\sqrt{(I^G)} = \langle x_1, \cdots, x_n \rangle$.

Proof: Let the assumptions be as above. It suffices to show $V(I^G) = V(\langle x_1, \cdots, x_n \rangle)$. Once this is accomplished we have $I(V(I^G)) = I(V(\langle x_1, \cdots, x_n \rangle))$. Then using the Hilbert Nullstellensatz we have $\sqrt{I^G} = I(V(I^G)) = I(V(\langle x_1, \cdots, x_n \rangle)) = \langle x_1, \cdots, x_n \rangle$.

Let $\vec{a} \in \mathbb{C}^n - \{0\}$. Define $G_{\vec{a}} = \{A \cdot \vec{a} : A \in G\}$. Since $G$ is finite then $G_{\vec{a}}$ is finite thus we can find a polynomial $f \in \mathbb{C}^n$ where $f(0) = 0$ and $f(A \cdot \vec{a}) = 1$ for all $A \in G$. Then consider $f^*(x) = \frac{1}{|G|} \sum_{A \in G} f(A \cdot \vec{x})$ which lies in $I^G$ since $f^*(0) = 0$, thus $\deg f^* > 0$. Then $f^*(a) = 1$ so we may say that $a \notin V(I^G)$. So $V(I^G) = V(\langle x_1, \cdots, x_n \rangle)$ and our desired result is proven.QED.

Now we can begin to develop our first algorithm for finding a set of fundamental invariants. This process will use Groebner bases, is algorithmic, and will be more efficient than the brute force effort already presented.

We will start off with the study of the Hilbert series. Hilbert series will help us avoid making frivolous checks and will save on computation time. We define the Hilbert series to be

$$\Phi_G(z) = \sum_{d=0}^{\infty} \dim(\mathbb{C}[\vec{x}]_d^G) z^d$$

where $\mathbb{C}[\vec{x}]_d^G$ is the set of all homogeneous invariants of degree $d$. Now the claim is

$$\Phi_G(z) = \frac{1}{|G|} \sum_{A \in G} \frac{1}{\det(I_n - zA)}.$$

The invariant of degree $d$ exists if and only if the coefficients of $z^d$ are nonzero since the coefficient is $\dim(\mathbb{C}[\vec{x}]_d^G)$. To prove our claim above we begin with a lemma. The following is the code for computing the Hilbert series in *Mathematica*

```
HilbertSer[G_, n_, S_]:=

Do[Print[Series[Together[(1/Length[G]) * Sum[1/CharacteristicPolynomial[L[i], z], {i, 1, Length[G]}]],

{z, 0, S}]], {k, 0, 0}]
```

**Lemma:** Let $G \subset GL_n(\mathbb{C})$ be a finite matrix group. Then the dimension of the invariant subspace $V^G = \{\vec{v} \in \mathbb{C}^n : A\vec{v} = \vec{v}$ for all $A \in G\}$ is equal to $\frac{1}{|G|} \sum_{A \in G} trace(A)$.

Proof: Let $P_G = \frac{1}{|G|} \sum_{A \in G} A$. This linear map is a projection onto the the invariant subspace $V^G$. Since the matrix $P_G$ defines a projection, we have $P_G = P_G^2$, which means that $P_G$ has only the eigenvalues 0 and 1. Therefore the rank of $P_G$ equals the multiplicity of its eigenvalues 1, and we fine $\dim(V^G) = rank(P_G) = trace(P_G) = \frac{1}{|G|} \sum_{A \in G} trace(A)$.QED.

This leads us to the proof of our claim above.

**Theorem:** The Hilbert series of the invariant ring $\mathbb{C}[\vec{x}]^G$ equals $\Phi_G(z) = \frac{1}{|G|}\sum_{A\in G}\frac{1}{\det(I_n - zA)}$.

Proof: We write $\mathbb{C}[\vec{x}]_d$ for the $\binom{n+d-1}{d}$-dimensional vector space of $d$-forms in $\mathbb{C}[\vec{x}]$. For every linear transformation $A \in G$ there is an induced linear transformation $A^d$ on the vector space $\mathbb{C}[\vec{x}]_d$. In this linear algebra notation $\mathbb{C}[\vec{x}]_d^G$ becomes precisely the invariant subspace of $\mathbb{C}[\vec{x}]_d$ with respect to the induced group $\{A^d : A \in G\}$ of $\binom{n+d-1}{d} \times \binom{n+d-1}{d}$-matrices.

In order to compute the trace of an induced transformation $A^d$, we identify the vector space $\mathbb{C}^n$ with its linear forms $\mathbb{C}[\vec{x}]_1$. Let $l_{A,1}, \cdots, l_{A,n} \in \mathbb{C}[\vec{x}]_1$. be the eigenvectors of $A = A^d$, and $\rho_{A,1}, \cdots, \rho_{A,n} \in \mathbb{C}$ denote the corresponding eigenvalues. Note that each matrix $A \in G$ is diagonalizable over $\mathbb{C}$ because it has finite order.

The eigenvectors of $A^d$ are precisely the $\binom{n+d-1}{d}$ $d$-forms $l_{A,1}^{d_1}\cdots l_{A,n}^{d_n}$ where $d_1 + \cdots + d_n = d$. The eigenvalues of $A^d$ are therefore the complex numbers $\rho_{A,1}^{d_1}\cdots\rho_{A,n}^{d_n}$ where $d_1 + \cdots + d_n = d$. Sine the trace of a linear transformation equals the sum of its eigenvalues, we have the equation

$$trace(A^d) \;=\; \sum_{d_1+\cdots+d_n=d} \rho_{A,1}^{d_1}\cdots\rho_{A,n}^{d_n}.$$

By the previous lemma, the dimension of the invariant subspace $\mathbb{C}[\vec{x}]_d^G$ equals the average of the traces of all group elements. Rewriting this dimension count in terms of the Hilbert series of the invariant ring, we get

$$\begin{aligned}
\Phi_G(z) &= \sum_{d=0}^{\infty}\frac{1}{|G|}\sum_{A\in G}\left(\sum_{d_1+\cdots d_n=d}\rho_{A,1}^{d_1}\cdots\rho_{A,n}^{d_n}\right)z^d \\
&= \frac{1}{|G|}\sum_{A\in G}\sum_{(d_1,\cdots,d_n)\in\mathbb{N}^n}\rho_{A,1}^{d_1}\cdots\rho_{A,n}^{d_n}z^{d_1+\cdots+d_n} \\
&= \frac{1}{|G|}\sum_{A\in G}\frac{1}{(1-z\rho_{A,1})\cdots(1-z\rho_{A,n})} \\
&= \frac{1}{|G|}\sum_{A\in G}\frac{1}{\det(I-zA)}.QED.
\end{aligned}$$

A **Graded Ring** is a the direct sum of additive subgroups: $S = S_0 \oplus S_1 \oplus S_2 \oplus \cdots$ such that $S_iS_j \subseteq S_{i+j}$

**Lemma:** Let $p_1, p_2, \cdots, p_m$ be algebraically independent elements of $\mathbb{C}[\vec{x}]$ which are homogeneous of degrees $d_1, \cdots, d_m$ repectively. Then the Hilbert series of the graded subring $R := \mathbb{C}[p_1, p_2, \cdots, p_m]$ equals

$$H(R, z) := \sum_{d=0}^{\infty}(\dim_{\mathbb{C}} R_d)z^d = \frac{1}{(1-z^{d_1})\cdots(1-z^{d_m})}.$$

Proof: Let $R_d$ denote the $\mathbb{C}$-vector space of degree $d$ elements in $R$. By the algebraic independence of $p_1, \cdots, p_m$ we have $\{p_1^{i_1} p_2^{i_2} \cdots p_m^{i_m} : i_1, \cdots, i_m \in \mathbb{N}$ and $i_1 d_1 + i_2 d_2 + \cdots + i_m d_m = d\}$ is a basis for $R_d$. This implies $\dim(R_d) = |A_d|$ where $A_d = \{(i_1, i_2, \cdots, i_m) \in \mathbb{N}^m : i_1 d_1 + \cdots + i_m d_m = d\}$.

So

$$\begin{aligned}
\sum_{d=0}^{\infty} (\dim_{\mathbb{C}} R_d) z^d &= \sum_{d=0}^{\infty} |A_d| z^d \\
&= \sum_{d=0}^{\infty} \sum_{(i_1, \cdots, i_m) \in A_d} z^d \\
&= \left( \sum_{i_1=0}^{\infty} z^{i_1 d_1} \right) \cdots \left( \sum_{i_m=0}^{\infty} z^{i_m d_m} \right) \\
&= \frac{1}{1 - z^{d_1}} \cdots \frac{1}{1 - z^{d_m}} \\
&= \frac{1}{(1 - z^{d_1}) \cdots (1 - z^{d_m})}
\end{aligned}$$

which is exactly what we wanted to show. QED.

If $R = \bigoplus_{i=1}^{t} \eta_i \mathbb{C}[\theta_1, \cdots, \theta_n]$ then by the previous lemma

$$R = \left( \bigoplus_{i=1}^{t} \eta_i \mathbb{C} \right) \oplus \left( \bigoplus_{(i_1, \cdots, i_n) \in \mathbb{N}^n - \{0\}} \bigoplus_{i=1}^{t} \eta_i \theta_1^{i_1} \cdots \theta_n^{i_n} \mathbb{C} \right)$$

we see that the Hilbert series for $R$ is

$$\frac{\sum_{i=1}^{t} z^{\deg \eta_i}}{\Pi_{j=1}^{n} (1 - z^{\deg \theta_j})}.$$

.

Then we conclude the following theorem immediately from the results above.

**Theorem:** $\Phi_G(z) \cdot \Pi_{j=1}^{n} (1 - z^{\deg \theta_j}) = \sum_{i=1}^{t} z^{\deg \eta_i}$.

The significance of the theorem above is that it allows us to know the degrees of the secondary invariants.

This section gives all the preliminary mathematical development in order to give and understand an algorithm for finding and computing the primary and secondary invariants.

## 5.5 Algorithms

*Algorithm 2:* Let $I := \langle f_1, \cdots, f_m \rangle$. Let $G$ be a Groebner basis of $\langle f_1, \cdots, f_m, gz - 1 \rangle$, where $z$ is a new ordered after all the previously existing variables. Then $g \in \sqrt{I}$ if and only if $1 \in G$.

Proof: Let $1 \in G$ then

$$1 \;=\; \sum_{i=1}^{s} p_i(x_1, \cdots, x_n, y) f_i + q(x_1, \cdots, x_n, y)(1 - yf)$$

for some $p_i, q \in K[x_1, \cdots, x_n, y]$. Let $y = \frac{1}{f(x_1, \cdots, x_n)}$ and the equation above implies

$$1 \;=\; \sum_{i=1}^{s} p_i(x_1, \cdots, x_n, \frac{1}{f}) f_i$$

Choose $m \in \mathbb{Z}^+$ so that when $f^m$ is multiplied by the equation above on both sides that the denominator is cleared. So we have

$$f^m \;=\; \sum_{i=1}^{s} h_i f_i \in I$$

for some $h_i \in K[x_1, \cdots, x_n]$. So $f \in \sqrt{I}$.

Let $f \in \sqrt{I}$. Then $f^m \in I \subset G$ for some $m \in \mathbb{Z}^+$. Note that $1 - yf \in G$ so

$$1 = y^m f^m + (1 - y^m f^m) = y^m f^m + (1 - yf)(1 + yf + \cdots + y^{m-1} f^{m-1}) \in G.$$

Therefore the algorithm with is implemented will work. QED.

```
Alg2[F_, g_, n_]:=

Do[{G = GroebnerBasis[Append[F, g * z - 1], Append[Array[a, n], z]],

If[Length[G]==1, p = 1, p = 0], Print[p]}, {i, 1, 1}]
```

*Algorithm 3:* Compute the Groebner basis $G$ for an ideal $I = \langle f_1, \cdots, f_m \rangle$. Then $\sqrt{I} = \langle x_1, \cdots, x_n \rangle$ if and only if a monomial of the form $x_i^{j_i}$ occurs among the initial monomials in $G$ for every $i$, for $1 \leq i \leq n$.

Proof: $\Rightarrow$ Suppose $\sqrt{I} = \langle x_1, \cdots, x_n \rangle$. Then we have $G = I \supset \{x_1^{j_1}, \cdots, x_n^{j_n}\}$ wich immediatly implies $LM(G) \supset \{x_1^{j_1}, \cdots, x_n^{j_n}\}$.

$\Leftarrow$ The key is to use the fact $f_1, \cdots, f_m$ are homogeneous and consider the monomial order as Lexicographic. First we must have $x_n^{j_n} \in G$ since $x_n^{j_n}$ is the leading term. Then we see that $x_n = 0$. Now suppose $x_{k+1} = \cdots = x_n = 0$. Then $x_k^{j_k} + \sum_{i_1+\cdots+i_n=j_k} a_k x_{k+1}^{i_{k+1}} \cdots x_n^{i_n}$ is a polynomial in $G$ by hypothesis for some $a_k \in \mathbb{C}$. But by our inductive hypothesis sum summand equals zero so $x_k^{j_k} = 0$. Therefore, $x_1 = \cdots = x_n = 0$ so $\sqrt{I} = \sqrt{G} = \langle x_1, \cdots, x_n \rangle$.QED.

```
Alg3[F_, n_]:=

Do[{LmF = {}, A = GroebnerBasis[F, Array[a, n]],

While[A ≠ {}, {f = Take[A, 1], A = Delete[A, 1],

LmF = Flatten[Append[LmF, Take[MonomialList[First[f]], 1]]]}], H = {},

While[LmF ≠ {}, {h = Take[LmF, 1], LmF = Delete[LmF, 1],

If[Length[Variables[h]] == 1, H = Flatten[Append[H, h]]]}],

If[Length[H] == n, Print[1], Print[0]]}, {i, 1, 1}]
```

*Algorithm 1:* Fix a monomial order $m_1 < m_2 < m_3 < \cdots$ which refines the partial order given by the total degree on the set of monomials of $\mathbb{C}[\vec{x}]$. Let $M = \langle x_1, \cdots, x_n \rangle$

0. Let $t := 1$ and $Q := \emptyset$.

1. While $m_t^* \notin Rad(\langle Q \rangle)$ let $t = t + 1$.

2. Let $Q := Q \cup \{m_t^*\}$.

3. If $|Q| < n$ and $\sqrt{\langle Q \rangle} \neq M$ then return to 1.

4. Either $Q$ is a primary set of generators $P$ or $Q$ can be modified to an algebraically independent set of primary generators $P$ of $n$ invariants with $\sqrt{\langle P \rangle} = M$. See [10] for more on this piece of the algorithm.

5. Find secondary invariants which are linearly independent module the ideal generated by $P$.

Proof: The termination of steps one and two is guaranteed by the lemma above. First note that $V(Q) = \{\vec{0}\}$. From this we see that $\{x_1^{d_1}, \cdots, x_n^{d_n}\} \subset \langle Q \rangle$ for some $d_1, \cdots, d_n \in \mathbb{Z}^+$. Then we can see

$$\mathbb{C}[\vec{x}] = \{r_1 a_1 + \cdots + r_m a_m : r_i \in \mathbb{C}[Q] \text{ and } a_j \in A\}$$

where $A = \{x_i^{j_i} : 0 \leq j_i < d_i \text{ for all } 0 \leq i \leq n\}$. A similar statement can be made about $\mathbb{C}[\vec{x}]^G$ thus $|Q| \geq n$. Then we are able to determine $P$ using Algorithm 3 which is stated below. Step 5 will work by taking Noether's bound into account.QED

For step 1, to check if $m_t^* \notin Rad(\langle Q\rangle)$ use *algorithm 2.* Moreover, the Hilbert series can be used to skip the power that will not be included.

For step 2, to check if $\sqrt{\langle Q\rangle} \neq M$ use *algorithm 3.*

For step 4, let $Q = \{q_1, \cdots, q_m\}$, $d_j = \deg(q_j)$ now find a $n \times m$-matrix matrix $(a_{i,j})$ where

$$\sqrt{\left\langle \sum_{j=1}^{m} a_{1,j} q_j^{d/d_j}, \cdots, \sum_{j=1}^{m} a_{n,j} q_j^{d/d_j} \right\rangle} \;=\; M.$$

This can be verified using algorithm 2.

For step 5, use $\Phi_G(z) \cdot \Pi_{j=1}^{n}(1 - z^{\deg \theta_j}) = \sum_{i=1}^{t} z^{\deg \eta_i}$ and the Reynolds operator to find the secondary invariants $\eta_1, \cdots, \eta_t$ which are linearly independent modulo the ideal generated by $P$. Let $GB$ be a Groebner basis of $\langle P\rangle$ then $\overline{\eta_j}^{GB} = 0$ if and only if $\eta_j$ is linearly independent modulo the ideal generated by $P$. It was discussed early on how to make this computation.

See the Appendix for code.

## 5.6 Algorithm Implementation

The algorithm is used to study cyclic and dihedral groups which can be represented by $2 \times 2$ matrices. Recall

$$D_{2m} \;=\; \left\langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right\rangle$$

where $\theta = \frac{2\pi}{m}$. Also the cyclic group of order $m$ is

$$C_m : \;=\; \left\langle \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right\rangle.$$

The generators of $\mathbb{C}[x,y]^{D_{2m}}$ and $\mathbb{C}[x,y]^{C_m}$ are interesting and, at first glance, unexpectedly related. For some values of $m$, for example 5 and 7, *Mathematica* has computational difficulty. The following tables presents the findings.

Table 5.1: Primary and secondary generators for $C_m$. (F means there was a computational fail)

| m | Primary | Secondary |
|---|---|---|
| 2 | $x^2, y^2$ | $1, xy$ |
| 3 | $x^2+y^2, x^3-3xy^2$ | 1 |
| 4 | $x^2+y^2, x^4+y^4$ | 1 |
| 5 | F | F |
| 6 | $x^2+y^2, 11x^6+15x^4y^2+45x^2y^4+9y^6$ | 1 |
| 7 | F | F |
| 8 | $x^2+y^2, 9x^8+28x^6y^2+70x^4y^4+28x^2y^6+9y^8$ | 1 |
| 9 | F | F |
| 10 | F | F |
| 11 | F | F |
| 12 | $x^2+y^2, 463x^{12}+2706x^{10}y^2+7425x^8y^4+8316x^6y^6+7425x^4y^8+2706x^2y^{10}+463y^{12}$ | 1 |

Table 5.2: Primary and secondary generators for $D_{2m}$.(F means there was a computational fail)

| m | Primary | Secondary |
|---|---|---|
| 2 | $x^2, y^2$ | 1 |
| 3 | $x^2+y^2, x^3-3xy^2$ | 1 |
| 4 | $x^2+y^2, x^4+y^4$ | 1 |
| 5 | F | F |
| 6 | $x^2+y^2, 11x^6+15x^4y^2+45x^2y^4+9y^6$ | 1 |
| 7 | F | F |
| 8 | $x^2+y^2, 9x^8+28x^6y^2+70x^4y^4+28x^2y^6+9y^8$ | 1 |
| 9 | F | F |
| 10 | F | F |
| 11 | F | F |
| 12 | $x^2+y^2, 463x^{12}+2706x^{10}y^2+7425x^8y^4+8316x^6y^6+7425x^4y^8+2706x^2y^{10}+463y^{12}$ | 1 |

An analysis of the tables 6.1 and 6.2 us that even though the dihedral groups and cyclic groups are structurally different and have different orders for each $m$ our generators are the same for all but one case. For cases three through twelve which worked this is easily explained by the fact that the flip, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which is a generator of $D_{2m}$, tells us that $y$ must be an even power. In each of the aforementioned cases $y$ has an even power. For the case $m = 2$ $y$ can not have and odd power so the secondary invariant is, in a sense, eliminated from $C_m$ to $D_{2m}$.

An another intriguing aspect is the radicals of each of the ideals generated by the primary invariants is equal to $\langle x, y \rangle$ and this is known immediately from the algorithm, and is true in general.

An interesting continuation would be to attempt to write this program in a more suitable computer algebra system and study $\mathbb{C}[x, y]^{D_{2m}}$ and $\mathbb{C}[x, y]^{C_m}$ and see if the patter continues for $m = 5, 7, 9, 10, 11$ and $m > 12$. The program may have implication issues when it comes to *mathematica* but the algorithm is implementable. Also, this rises the question if we have a set of generators can we always find which groups will give us these generators? We have satisfied one of the goals and have given the solution to a special case of Hilbert's Fourteenth Problem.

# Chapter 6

# Generalized Groebner Basis Theory and The Straightening Law

## 6.1 Generalized Groebner Basis Theory

Generalized Groebner basis theory, as it pertains to us, can be found in [9]. So far we have seen the notion of a Groebner basis in a the setting of a polynomial ring over a field. This chapter is a consequence of a more general theory. Groebner basis theory can be applied to non-commutative rings, infinite dimensional algebras, and, in our case, bracket rings. The initial goal is to introduce Groebner bases in a more general setting and to state its equivalence to the case that has been used up to this point of the thesis. The second goal is to prove the so called Straightening Law. The Straightening Law gives us a $\mathbb{C}$-vector basis for the Bracket ring which is a key tool in proving the Fundamental Theorem of Invariant Theory.

Consider $\mathbb{C}[x_1, \cdots, x_n]$ with the ordering $x_1 < \cdots < x_n$. Now consider the set $\mathbb{M} = \{x_1^{i_1} \cdots x_n^{i_n} : \text{for all } i_j \geq 0\}$. A total ordering on $\mathbb{M}$ is said to be **admissible** if $1 \leq m$ for all $m \in \mathbb{M}$, and for all $p, q, r \in \mathbb{M}$ we have $p < q$ implies $p \cdot r < q \cdot r$. The monomial orders presented earlier are all examples of admissible orders.

Let $F \subset \mathbb{C}[x_1, \cdots, x_n]$ and $g, h \in \mathbb{C}[x_1, \cdots, x_n]$. If $h = g - b \cdot u \cdot f$ for some $f \in F, b \in \mathbb{C}$, and $u \in \mathbb{M}$ where $b \cdot u \cdot LM(f)$ is a monomial of $g$ then $g$ is defined to **reduce** to $h$, denoted $g \to_F h$. Moreover, $h$ is in **reduced form** provided there does not exist $h'$ where $h \to_F h'$. If there exists a sequence of reduction

$$g \to_F h_1 \to_F \cdots \to_F h_k \to_F h$$

where $h$ is in reduced form then $h$ is the **normal form** of $g$. Now it is important to note that $h$ is not necessarily unique. In the case where $F$ is finite the normal form of $g$ is simply the remainder upon the division of $F$.

$F$ is a **"Generalized" Groebner basis** of $\langle F \rangle$ if $g \in \mathbb{C}[x_1, \cdots, x_n]$ has a unique normal form modulo $F$. The thoerem below shows a Groebner basis is a "Generalized" Groebner basis.

**Theorem:** Let $F = \{f_1, \cdots, f_j\}$. Then The following are equivalent:

i) $F$ is a "Generalized" Groebner basis for $\langle F \rangle$

ii) For all $f, g \in K[x_1, \cdots, x_n]$, $f + \langle F \rangle = g + \langle F \rangle$ if and only if the normal forms, reduced using $F$, of $f$ and $g$ are equal.

(iii) $LM(\langle F \rangle) = LM(F)$

Proof: (i)$\Leftrightarrow$(ii) The result follows immediately from definition.

(iii) $\Rightarrow$(i) Let $LM(\langle F \rangle) = LM(F)$. Let $f, g \in \mathbb{C}[x_1, \cdots, x_n]$ with $f + \langle F \rangle = g + \langle F \rangle$ and $f, g$ are both in reduced form. First note that $LM(f), LM(g) \notin LM(F)$, else either $f$ or $g$ would be reducible.

Now by definition $f - g \in \langle F \rangle$ which implies $LM(f - g) \in LM(\langle F \rangle) = LM(F)$. Thus we can conclude that $LM(f) = LM(g)$ and $LC(f) = LC(g)$.

Let $f' = f - LC(f) \cdot LM(f)$ and $g' = g - LC(g) \cdot LM(g)$. If $f'$ were reducible then $LM(f_i)$, for some $i$, would divide some monomial of $f'$, and consequently $f$. So $f$ is irreducible. By analogous reasoning $g'$ is irreducible. So we may conclude, after repeating the steps above, that $f = g$. Therefore we have $F$ is a Generalized Groebner basis.

(i)$\Rightarrow$(iii) can be found in [9]. QED.

## 6.2 Straightening Law in Terms of Groebner Bases

This final section is drawing connections between topics covered in [7] and [6]. Let $X = (x_{i,j})$ be a $n \times d$-matrix. A maximal minor of $X$ is the determinate of the $d \times d$ matrix using $d$ rows from $X$. We define all the maximal minors of $X$ to be the *Plücker coordinates* of X. The **Ring of Plücker Coordinates** is the polynomial ring generated by the Plücker Coordinates.

**First Fundamental Theorem of Invariant Theory:** Let $\mathbb{C}[x_{i,j}]^{SL_d(\mathbb{C})} = \{f \in \mathbb{C}[x_{i,j}] :$

$f(X \cdot A) = f(X)$ for all $A \in SL_d(\mathbb{C})\}$. The ring $\mathbb{C}[x_{i,j}]^{SL_d(\mathbb{C})}$ equals the ring of Plucker coordinates of $X$.

**Example:** Let $P(x_{i,j}) \in \mathbb{C}[x_{i,j}]$ be a Plücker coordinate. Now, by defintion, $P(x_{i,j}) = \det M$ where $M$ is some maximal minor of $(x_{i,j})$. Let $A \in SL_d(\mathbb{C})$ then we see that $P(A \cdot (x_{i,j})) = \det(A \cdot M) = \det A \cdot \det M = \det M = P(x_{i,j})$. So $P(x_{i,j}) \in \mathbb{C}[x_{i,j}]^{SL_d(\mathbb{C})}$. So we can conclude that the Plücker coordinates are invariants.

To help motivate the technique and theory used to prove the statement above we'll consider an analogous case in a setting made familiar in this paper.

Let $K[x_1, \cdots, x_n]^G = K[f_1, \cdots, f_m]$ where $G$ is a finite subset of $GL_n(K)$. Define $\phi : K[y_1, \cdots, y_m] \to K[x_1, \cdots, x_n]^G$ by $\phi(g) = g(f_1, \cdots, f_m)$. $\phi$ is well-defined by construction. Let $g, h \in K[y_1, \cdots, y_m]$ then

$$\begin{aligned} \phi(g+h) &= (g+h)(f_1, \cdots, f_m) \\ &= g(f_1, \cdots, f_m) + h(f_1, \cdots, f_m) \\ &= \phi(g) + \phi(h). \end{aligned}$$

which establishes $\phi$ as a homomorphism. $\phi$ is an onto map since $K[x_1, \cdots, x_n]^G = K[f_1, \cdots, f_m]$. Now it is enough to say that $K[y_1, \cdots, y_m]/\ker\phi \cong K[x_1, \cdots, x_n]^G$.

Now by definition $\ker\phi = \{h \in K[y_1, \cdots, y_m] : h(f_1, \cdots, f_m) = 0 \text{ in } K[x_1, \cdots, x_n]\}$. $\ker\phi$ establishes the relationships between elements of $K[y_1, \cdots, y_m]$. Let $g_1, g_2 \in K[y_1, \cdots, y_m]$ then $g_1(f_1, \cdots, f_m) = g_2(f_1, \cdots, f_m)$ if and only if $(g_1 - g_2)(f_1, \cdots, f_m) = 0$. Thus $g_1 - g_2 \in \ker\phi$.

Let $\Lambda(n,d) = \{X_{\lambda_1, \cdots, \lambda_d} : 1 \leq \lambda_1 < \lambda_2 < \cdots < \lambda_d \leq n\}$. Let $\pi \in S_d$ then $X_{\lambda_{\pi(1)}, \cdots, \lambda_{\pi(d)}} = \text{sign}(\pi) \cdot X_{\lambda_1, \cdots, \lambda_d}$.

Let $\theta_{n,d} : \mathbb{C}[\Lambda(n,d)] \to \mathbb{C}[x_{i,j}]$ be defined by

$$X_{\lambda_1, \cdots, \lambda_d} \quad \mapsto \quad \det \begin{pmatrix} x_{\lambda_1,1} & x_{\lambda_1,2} & \cdots & x_{\lambda_1,d} \\ x_{\lambda_2,1} & x_{\lambda_2,2} & \cdots & x_{\lambda_2,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{\lambda_d,1} & x_{\lambda_d,2} & \cdots & x_{\lambda_d,d} \end{pmatrix}.$$

This the image of $\theta_{n,d}$ is the ring of Plucker coordinates. By properties of determinant we have $\theta_{n,d}$ is a homomorphism. In addition we see that $\mathbb{C}[\Lambda(n,d)]/\ker\theta_{n,d} \cong Im(\theta_{n,d})$.

Let's discuss how to use Groebner bases to find the generators of $\ker\phi$ for an arbitrary ring homomorphism $\phi$. To see how this can be done we present the following theorem.

**Theorem:** Let $K[x_1,\cdots,x_n]^G = K[f_1,\cdots,f_m]$, and consider the ideal $J = \langle f_1 - y_1,\cdots,f_m - y_m\rangle \subset K[x_1,\cdots,x_n,y_1,\cdots,y_m]$.

Then $\ker\phi$ is the $n$-th elimination ideal of $J$, and so $\ker\phi = J \cap K[y_1,\cdots,y_m]$.

Proof: Let $p \in J$. Note if we substitute $y_i \mapsto f_i$ for all $i$ then $J = \langle 0\rangle$. Therefore $p(x_1,\cdots,x_n,f_1,\cdots,f_m) = 0$.

Let $p(x_1,\cdots,x_n,f_1,\cdots,f_m) = 0$ in $K[x_1,\cdots,x_n]$. Now let $y_i = f_i - (f_i - y_i)$ for all $i$. Consider the following monomial

$$\begin{aligned} x_1^{a_1}\cdots x_n^{a_n} y_1^{b_1}\cdots y_m^{b_m} &= x_1^{a_1}\cdots x_n^{a_n}(f_1 - (f_1 - y_1))^{b_1}\cdots(f_m - (f_m - y_m))^{b_m} \\ &= x_1^{a_1}\cdots x_n^{a_n}\left[y_1^{b_1}\cdots y_m^{b_m} + B_1(f_1 - y_1) + \cdots + B_m(f_m - y_m)\right] \end{aligned}$$

for some $B_1,\cdots,B_m \in K[x_1,\cdots,x_n,y_1,\cdots,y_m]$. Multiplying by an appropriate constant and adding over the exponents appearing in $p$ we arrive at

$$p(x_1,\cdots,x_n,y_1,\cdots,y_m) = p(x_1,\cdots,x_n,f_1,\cdots,f_m) + C_1(f_1 - y_1) + \cdots + C_m(f_m - y_m)$$

For some $C_1,\cdots,C_m \in K[x_1,\cdots,x_n,y_1,\cdots,y_m]$. Since $p(x_1,\cdots,x_n,f_1,\cdots,f_m) = 0$ we have

$$p(x_1,\cdots,x_n,y_1,\cdots,y_m) = C_1(f_1 - y_1) + \cdots + C_m(f_m - y_m) \in J.$$

Now we have established $p \in J$ if and only if $p(x_1,\cdots,x_n,f_1,\cdots,f_m) = 0$. Taking intersections we now have $p \in J \cap K[y_1,\cdots,y_m]$ if and only if $p(f_1,\cdots,f_n) = 0$ in $K[x_1,\cdots,x_n]$. Therefore $\ker\phi = J \cap K[y_1,\cdots,y_m]$.QED.

**Corollary:** Fix a monomial order in $K[x_1,\cdots,x_n,y_1,\cdots,y_m]$ where any monomial involving $x_1,\cdots,x_n$ is great than all monomials in $K[y_1,\cdots,y_m]$ and let $G$ be a Groebner basis of $J$. Then $G \cap K[y_1,\cdots,y_m]$ is a Groebner basis for $I$ in the induced monomial order induced on $K[y_1,\cdots,y_m]$.

Proof: The result follows immediately from work already presented on elimination theory.

The two results above give us an algorithm for finding the generators of $\ker\phi$. Now how do we find the generators of $\ker\theta_{n,d}$.

Let us identify $X_{\lambda_1,\cdots,\lambda_d}$ with $[\lambda_1\cdots\lambda_d] = \lambda$. We will now introduce the van der Waerden syzygy with it's associated notation. First let $\lambda \in \Lambda(n,d)$ then $\lambda^c \in \Lambda(n,n-d)$ with $[\lambda] \cup \lambda^c = \{1,\cdots,n\}$. Let $(\lambda,\lambda^c) = sign(\pi)$ where $\pi$ is the permutation where $\lambda_i \mapsto i$ for $i = 1,\cdots,d$ and $\lambda_j^c \mapsto d + j$ for $j = 1,\cdots,n-d$.

Let $s \in \{1, \cdots, d\}$, $\alpha \in \Lambda(n, s-1)$, $\beta\Lambda(n, d+1)$, and $\gamma \in \Lambda(n, d-s)$. Then the Van der Waerden syzygy is

$$[[\alpha\bar{\beta}\gamma]] \quad = \quad \sum_{\tau\in\Lambda(d+1,s)} (\tau, \tau^c) \cdot [\alpha_1 \cdots \alpha_{s-1}\beta_{\tau_1^c} \cdots \beta_{\tau_{d+1-s}^c}] \cdot [\beta_{\tau_1} \cdots \beta_{\tau_{d+1-s}}\gamma_1 \cdots \gamma_{d-s}].$$

For an example of what these look like let $d = 5, s = 4$ and $n$ large enough and considering

$$\begin{aligned} \alpha &= [\alpha_1\alpha_2\alpha_3] \in \Lambda(n, 3) \\ \beta &= [\beta_1\beta_2\beta_3\beta_4\beta_5\beta_6] \in \Lambda[n, 6] \\ \gamma &= [\gamma_1] \in \Lambda(n, 1). \end{aligned}$$

We then have

$$\begin{aligned} [[\alpha\bar{\beta}\gamma]] &= [\alpha_1\alpha_2\alpha_3\bar{\beta}_1\bar{\beta}_2\bar{\beta}_3\bar{\beta}_4\bar{\beta}_5\bar{\beta}_6\gamma_1] \\ &= (\tau_1, \tau_1^c)[\alpha_1\alpha_2\alpha_3\beta_5\beta_6][\beta_1\beta_2\beta_3\beta_4\gamma_1] \\ &+ (\tau_2, \tau_2^c)[\alpha_1\alpha_2\alpha_3\beta_4\beta_6][\beta_1\beta_2\beta_3\beta_5\gamma_1] \\ &+ (\tau_3, \tau_3^c)[\alpha_1\alpha_2\alpha_3\beta_4\beta_5][\beta_1\beta_2\beta_3\beta_6\gamma_1] \\ &+ (\tau_4, \tau_4^c)[\alpha_1\alpha_2\alpha_3\beta_3\beta_6][\beta_1\beta_2\beta_4\beta_5\gamma_1] \\ &+ (\tau_5, \tau_5^c)[\alpha_1\alpha_2\alpha_3\beta_3\beta_5][\beta_1\beta_2\beta_4\beta_6\gamma_1] \\ &+ (\tau_6, \tau_6^c)[\alpha_1\alpha_2\alpha_3\beta_3\beta_4][\beta_1\beta_2\beta_5\beta_6\gamma_1] \\ &+ (\tau_7, \tau_7^c)[\alpha_1\alpha_2\alpha_3\beta_2\beta_6][\beta_1\beta_3\beta_4\beta_5\gamma_1] \\ &+ (\tau_8, \tau_8^c)[\alpha_1\alpha_2\alpha_3\beta_2\beta_5][\beta_1\beta_3\beta_4\beta_6\gamma_1] \\ &+ (\tau_9, \tau_9^c)[\alpha_1\alpha_2\alpha_3\beta_2\beta_4][\beta_1\beta_3\beta_5\beta_6\gamma_1] \\ &+ (\tau_{10}, \tau_{10}^c)[\alpha_1\alpha_2\alpha_3\beta_2\beta_3][\beta_1\beta_4\beta_5\beta_6\gamma_1] \\ &+ (\tau_{11}, \tau_{11}^c)[\alpha_1\alpha_2\alpha_3\beta_1\beta_6][\beta_2\beta_3\beta_4\beta_5\gamma_1] \\ &+ (\tau_{12}, \tau_{12}^c)[\alpha_1\alpha_2\alpha_3\beta_1\beta_5][\beta_2\beta_3\beta_4\beta_6\gamma_1] \\ &+ (\tau_{13}, \tau_{13}^c)[\alpha_1\alpha_2\alpha_3\beta_1\beta_4][\beta_2\beta_3\beta_5\beta_6\gamma_1] \\ &+ (\tau_{14}, \tau_{14}^c)[\alpha_1\alpha_2\alpha_3\beta_1\beta_3][\beta_2\beta_4\beta_5\beta_6\gamma_1] \\ &+ (\tau_{15}, \tau_{15}^c)[\alpha_1\alpha_2\alpha_3\beta_1\beta_2][\beta_3\beta_4\beta_5\beta_6\gamma_1] \end{aligned}$$

Just note that $\binom{6}{4} = 15$ is the number of brackets being summed. Now let's compute $(\tau_1, \tau_1^c)$. First note that $\tau_1 = [1, 2, 3, 4]$ so $\tau_1^c = [5, 6]$. So the associated permutation $\pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = (1)$. Thus $(\tau_1, \tau_1^c) = 1$ since $\pi_1$ is the identity.

Let's compute $(\tau_2, \tau_2^c)$. First note that $\tau_2 = [1, 2, 3, 5]$ so $\tau_1^c = [4, 6]$. So the associated permutation $\pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 4 & 6 \end{pmatrix} = (4, 5)$. Thus $(\tau_2, \tau_2^c) = -1$ since $\pi_2$ is an odd permuation.

For a final example let's consider $(\tau_3, \tau_3^c)$. First note that $\tau_3 = [1,2,3,6]$ and $\tau_3 = [4,5]$. So $\pi_3 = (6,4,5)$ which is even. So $(\tau_3, \tau_3^c) = 1$.

So we have established what the van der Waerden syzygiess are and an example of what they look like. A van der Waerden syzygy is called a straightening syzygy provided $\alpha_{s-1} < \beta_{s+1}$ and $\beta_s < \gamma_1$. Let

$$S_{n,d} = \{[[\alpha\bar{\beta}\gamma]] : \alpha_{s-1} < \beta_{s+1} \text{ and } \beta_s < \gamma_1\}$$

The claim is $S_{n,d}$ is a Groebner basis for $I_{n,d}$. This result is a stronger condition than the *Second Fundamental Theorem of Invariant Theory.*

**Lemma:** Let $I$ be any ideal and $<$ be any monomial order on $\mathbb{C}[x_1, \cdots, x_n]$. The set $S := \{m + I : m \notin LM(I)\}$ is a $\mathbb{C}$-vector space basis for the ring $\frac{\mathbb{C}[x_1,\cdots,x_n]}{I}$.

Proof: Let $g + I \in \frac{\mathbb{C}[x_1,\cdots,x_n]}{I}$ where $g = \sum_{i=1}^{k_1} c_i x^{\alpha_i} + \sum_{i=1}^{k_1} d_i x^{\beta_i}$ where $x^{\alpha_i} + I \in S$ and $x^{\beta_j} + I \notin S$. So we have

$$\begin{aligned} g + I &= \left(\sum_{i=1}^{k_1} c_i x^{\alpha_i} + \sum_{i=1}^{k_2} d_i x^{\beta_i}\right) + I \\ &= \sum_{i=1}^{k_1} (c_i x^{\alpha_i} + I) + \sum_{i=1}^{k_2} \left(d_i x^{\beta_i} + I\right) \\ &= \sum_{i=1}^{k_1} (c_i x^{\alpha_i} + I) + \sum_{i=1}^{k_2} (0 + I) \\ &= \sum_{i=1}^{k_1} (c_i x^{\alpha_i} + I) \end{aligned}$$

So every element of $\frac{\mathbb{C}[x_1,\cdots,x_n]}{I}$ can be represented from elements of $S$.

Let $0 + I = \sum_{i=1}^{k_1} (c_i x^{\alpha_i} + I)$. Then $\sum_{i=1}^{k_1} c_i x^{\alpha_i} \in I$. Now every polynomial in $I$ must have a monomial $m$, by definition, where $m + I \notin S$. This implies $c_i = 0$ for all $i$ and hence unique representations from elements in $S$. Therefore we have established $S$ is a vector space basis. QED.

Define $<_T$ to be the ordering on $\Lambda(n,d)$ where $[\lambda] <_T [\mu]$ provided there exists $m$, $1 \leq m \leq d$, such that $\lambda_j = \mu_j$ for $1 \leq j \leq m-1$ and $\lambda_m < \mu_m$. $<_T$ ordering the elements of $\Lambda(n,d)$ lexicographically.

$<_T$ can be generalized to an order on $\mathbb{C}[\Lambda(n,d)]$ and we'll call $<_T$ the tableaux order. Define $T = [\lambda^1] \cdots [\lambda^k]$ where $[\lambda_i] = [\lambda_1^i \cdots \lambda_d^i]$, $[\lambda^1], \cdots, [\lambda^k] \in \Lambda(n,d)$, and $[\lambda^1] <_T \cdots <_T [\lambda^k]$. We will say $T$ is standard provided $\lambda_s^1 \leq \cdots \lambda_s^k$ for all $s$. Otherwise $T$ is nonstandard.

Let $T$ be as above and $T' = [\mu^1] \cdot [\mu^l]$ where $k \leq l$. Then we define $T < T'$ if either $k < l$ or there exists $r \in \{1, 2, \cdots, k\}$ and $s \in \{1, \cdots, d\}$ such that $\lambda_j^i = \mu_j^i$ for all $i$ and $j$ such that $i < r$ or $i = r$ and $j < s$, and $\lambda_s^r < \mu_s^r$.

**Lemma:** The polynomials $[[\alpha\bar{\beta}\gamma]]$ are contained in the ideal $I_{n,d}$.

The ideal of this proof is to show that $[[\alpha\bar{\beta}\gamma]$ in the the kernal of $\ker\theta_{n,d}$.QED

The significance of the Lemma is to show that we have a subset of $I_{n,d}$ and the Groebner basis for $I_{n,d}$ are elements of the form $[[\alpha\bar{\beta}\gamma]]$ with an additional property that $\alpha_{s-1} < \beta_{s+1}$ and $\beta_s < \gamma_1$. This is set is denoted by $S_{n,d}$ and an element is called a *straightening sygyzy.* The following theorem is the formal statement.

**Theorem:** The set $S_{n,d}$ of straightening syzygies forms a Groebner basis for the syzygy ideal $I_{n,d}$ with respect to the tableaux order.

Proof: Follows from the previous Lemma by the definition of Groebner basis. It is shown that the set $M$ of monomial idea generated by the initial tableaux of the elements of $S_{n,d}$ contains $LM(I_{n,d})$. It is shown that all nonstandard tableau are elements of $M$ and then it is proven that each monomial in $LM(I_{n,d})$ is a nonstardard tableau. QED.

The following theorem makes an important connection with the notion of leading monomials and standard tableaux. It also allows us to know the form of a leading monomial and is key to proving the straightening algorithm.

**Theorem:** $T$ is a standard tableaux if and only if $T \notin LM(I_{n,d})$.

Proof: Follows from the same line of reasoning as the proof above.QED.

We have a Groebner basis of $I_{n,d}$ we can state the following result. The importance is we can write any bracket in terms of standard tableaux uniquely and as a result simplifies many proofs to a single case.

**Corollary (Straightening Law):** The standard tableaux form a $\mathbb{C}$-vector space basis for the Plucker ring.

Proof: Follows immediately from the previous Lemmas and Theorems.QED.

The corollary is an essential piece of invariant theory and is a key piece of machinery, which was developed using Groebner basis theory, when proving the First Fundamental Theorem of invariant theory.

This section presents, not only, higher level Invariant theory, but also an example of a generalized notion of a Groebner basis.

# Chapter 7

# Appendix A

## 7.1 Code for *algorithm 1*

We begin by declaring the group. The L[i] represent the group elements and G is the group so this first part finds the primary invariants.

```
Clear[“Global`*”]

L[1] = {{1, 0}, {0, 1}};

L[2] = {{−1, 0}, {0, −1}};

L[3] = {{0, −1}, {1, 1}};

L[4] = {{−1, −1}, {1, 0}};

L[5] = {{0, 1}, {−1, −1}};

L[6] = {{1, 1}, {−1, 0}};

G = {L[1], L[2], L[3], L[4], L[5], L[6]};
```

Now we declare what n, LG, P, and B[i] are for future use.

```
n = 2;

LG = Length[G];

P = {};

Do[B[i] = L[i], {i, 1, Length[G]}]
```

Now we are constructing P to equal integer partitions of i into at most n integers.

```
Do[P = Union[IntegerPartitions[i, n], P], {i, 1, LG}];
```

This piece of the code is parameterizing the integers partitions found above by s[i] into the list S. This code is also permuting the elements so we have, for example, {0,1} and {0,1}.

```
S = {};
```

Now we are taking the integer partitions of S and using them as weight vectors and creating a long polynomial composed of all the monomials.

```
M = 0;
```

L is the list of monomials in the polynomial above in reverse graded lexicographic order.

```
L = Flatten[MonomialList[M, Array[x, n], “NegativeDegreeLexicographic”]];
```

In this piece of code we are parameterize the monomials of L by m[i]

```
i = 1;
```

Here we are parameterizing the entries of the matrices after multiplication.

```
U = {};
k = 1;

While[G ≠ {}, {U = Flatten[Take[G, 1], 1], i = 1,
While[U ≠ {}, {U1 = Flatten[Take[U, 1], 1], j = 1,
While[U1 ≠ {}, {w[k][i][j] = Flatten[Take[U1, 1], 1], U1 = Drop[U1, 1],
j = j + 1}], U = Drop[U, 1], i = i + 1}], G = Drop[G, 1], k = k + 1}];
```

```
Do[Do[x[j][i] = Sum[w[j][i][j1] * x[j1], {j1, 1, n}],

{i, 1, n}], {j, 1, LG}];
```

This is the Reynold's operator using the above construction and code. Note that the reynolds operator only works in two variables.

```
F = {0};

t = 0;


While[Length[F] == 1,

{t = t + 1, Do[mo[j][t] = Flatten[m[t]/.{x[1] → x[j][1], x[2] → x[j][2]}],

{j, 1, LG}],

g[t] = (1/LG) * Sum[mo[j][t], {j, 1, LG}], F = Union[F, g[t]]}];
```

This piece of code is the algorithm above. Algorithm 2 and Algorithm 3, presented as well, are also included in this piece of the code.

```
p = 1;

q = 1;


While[q==1, {While[p!=1;


{t = t + 1, Do[mo[j][t] = Flatten[m[t]/.{x[1] → x[j][1], x[2] → x[j][2]}],

{j, 1, LG}],

g[t] = (1/LG) * Sum[mo[j][t], {j, 1, LG}],

g1[t] = (z/LG) * Sum[mo[j][t], {j, 1, LG}],

G = GroebnerBasis[Append[F, Part[g1[t], 1] - 1], Append[Array[x, n], z]],

If[Length[G]==1, p = 0, p = 1], If[p==1, F = Union[F, {Part[g[t], 1]}]]}],
LmF = {}, A = GroebnerBasis[F, Array[x, n]],
While[A ≠ {}, {f = Take[A, 1], A = Drop[A, 1],
LmF = Flatten[Append[LmF, Take[MonomialList[First[f]], 1]]]}],
H = {}, While[LmF ≠ {}, {h = Take[LmF, 1], LmF = Delete[LmF, 1],
If[Length[Variables[h]] == 1, H = Flatten[Append[H, h]]]}],
If[Length[H] == n, q = 0, q = 1]


}];
```

```
F1 = Drop[F, 1];
Print[F1]
```

$\left\{\frac{1}{6}\left(2x[1]^2+(-x[1]-x[2])^2+2x[2]^2+(x[1]+x[2])^2\right),\frac{1}{6}\left(2x[1]^6+(-x[1]-x[2])^6+2x[2]^6+(x[1]+x[2])^6\right)\right.$

The output below is the expected output through step 3 of Algorithm 1.

Now we cover the case when the above returns a list with 3 or more elements. First we are going to choose a matrix of size 2 by the size of the list above where R[i] is the i row.

```
R[1] = {1, 0};
R[2] = {0, 1};
```

This next piece is parameterizing the degrees of the polynomials found above by d[i].

```
Do[
{y[j] =
Part[Part[Flatten[Part[CoefficientRules[Take[F1, 1], Array[x, n]]], 1], 1],
1], d[j] = 0,
While[y[j] ≠ {}, {d[j] = Take[y[j], 1] + d[j], y[j] = Drop[y[j], 1]}]},
{j, 1, Length[F1]}];
```

This is find the least common multiple of the degrees.

```
Do[{c = d[1], c = LCM[c, d[j]]}, {j, 1, Length[F1]}];
```

This is taking each of the polynomials found above and raising them to the c/d[j] power.

```
F2 = {};
```

Now we multiply by our previous choice of matrix.

```
F1 = {R[1].F1, R[2].F1};
```

The final part uses algorithm 3 to decided whether F1 has the trivial ideal as its radical.

```
LmF = {};
```

```
A = GroebnerBasis[F, Array[x, n]];
While[A ≠ {}, {f = Take[A, 1], A = Delete[A, 1],
LmF = Flatten[Append[LmF, Take[MonomialList[First[f]], 1]]]}];
```

```
H = {};
```

1

$$\left\{\frac{1}{6}\left(2x[1]^2+(-x[1]-x[2])^2+2x[2]^2+(x[1]+x[2])^2\right),\frac{1}{6}\left(2x[1]^6+(-x[1]-x[2])^6+2x[2]^6+(x[1]+x[2])^6\right)$$

The output above is 1 if our matrix worked. The output is 0 if the matrix did not work. We also get our new list of polynomials as an output.

Now our goal is to find the secondary invariants. The first part parameterizes the degrees of the primary invariants.

```
Do[
{y[j] =
Part[Part[Flatten[Part[CoefficientRules[Take[F1, 1], Array[x, n]]], 1], 1],
1], d[j] = 0,
While[y[j] ≠ {}, {d[j] = Take[y[j], 1] + d[j], y[j] = Drop[y[j], 1]}]},
{j, 1, Length[F1]}];
```

Now we are finding the polynomial described in step 5 of algorithm 1.

```
Pr = Expand[Flatten[Product[1 − z^{d[i]}, {i, 1, Length[F1]}]]];


HIl =
Together[(1/Length[G]) ∗ Sum[Pr /CharacteristicPolynomial[B[i], z],
{i, 1, Length[G]}]];
```

Now we are find the monomial list for the polynomial above.

```
Mon = Flatten[MonomialList[HIl, z]];
```

Now we are parameterizing the degrees of the monomials in the polynomial above.

```
a = 1;


SecD = {};
```

Now we are finding the integer partitions for the degree of the monomials above with at most n integers.

```
P1 = {};
```

This piece of the code is parameterizing the integers partitions found above by r[i] into the list S1. This code is also permuting the elements so we have, for example, {0,1} and {0,1}.

```
P2 = P1;
```

```
S1 = {};
```

Now we are taking the interger partitions of S1 and using them as weight vectors and creating a long polynomial composed of all the monomials.

```
M1 = 0;
```

L1 is the list of monomials in the polynomial above in reverse graded lexicographic order.

```
L1 = Flatten[MonomialList[M1, Array[x, n], “NegativeDegreeLexicographic”]];
```

In this piece of code we are parameterizing the monomials of L1 by m[i2]

```
i2 = 1;
```

Here we are are parameterizing the entries of the matrices after multiplication.

```
U2 = {};
k = 1;
```

```
While[G1 ≠ {}, {U2 = Flatten[Take[G1, 1], 1], i2 = 1,
While[U2 ≠ {}, {U3 = Flatten[Take[U2, 1], 1], j = 1,
While[U3 ≠ {}, {w[k][i2][j] = Flatten[Take[U3, 1], 1], U3 = Drop[U3, 1],
j = j + 1}], U2 = Drop[U2, 1], i2 = i2 + 1}], G1 = Drop[G1, 1], k = k + 1}];
```

```
Do[Do[x[j][i] = Sum[w[j][i][j1] * x[j1], {j1, 1, n}],
{i, 1, n}], {j, 1, LG}];
```

This is the Reynold's operator using the above construction and code. Note that the reynolds operator only works in two variables.

```
F = {0};

t = 0;


Do[{Do[mo1[j][t] = Flatten[m[t]/.{x[1] → x[j][1], x[2] → x[j][2]}], {j, 1, LG}],

g2[t] = (1/LG) ∗ Sum[mo1[j][t], {j, 1, LG}], F = Union[F, g[t]]},

{t, 1, Length[P2]}];
```

The last piece of code reduces the secondary invariants modulo the Groebner basis of the primary set of invariants to remove redundancy.

```
GB = GroebnerBasis[F1, Array[x, n]];


F2 = {};


Do[{{{Q, R[t]}} = PolynomialReduce[g2[t], GB, Array[x, n]],

F2 = Union[F2, {R[t]}]}, {t, 1, Length[P2]}];
```

```
{0, 1}
```

The output above is the secondary set of invariants.

# Chapter 8

# Appendix B

## 8.1 Explanations of Certain *Mathematica* Commands

Explanations and examples of *Mathematica* commands used. The explanations and more examples can be found in the documentation center in *Mathematica* under help. The examples I give are very specific but are the most relevant to how I used the commands.

*Clear["Global*"]* clears all the variables globally.

*Length[G]* outputs the length of $G$. For example if $G = \{1, 1, 7\}$ then the output would be 3.

*IntegerPartitions[i,n]* returns all the partitions of $i$ in at most $n$ integers. For example if $i = 3$ and $n = 5$ the output would be $\{\{3\}, \{2, 1\}, \{1, 1, 1\}\}$.

*Flatten[A]* takes a layer of parenthesis off $A$. If $A = \{\{1, 3, 4\}\}$ then the output would be $\{1, 3, 4\}$.

*MonomialList[f]* gives all the monomials of the polynomial $f$. If $f = x^2y + yz + z^4$ then the output would be $\{x^2y, yz, z^4\}$.

*Take[A,1]* returns the first element in the list $A$. So if $A = \{5, 4, 6\}$ then the output would be 5.

*Drop[A,1]* deletes the first element of a list. So if $A = \{5, 4, 6\}$ then the output would be $\{4, 6\}$.

*Append[A, x]* includes $x$ at the end of the list of $A$. If $A = \{a, b, c, d\}$ then the output would be $\{a, b, c, d, x\}$.

*CoefficentRules[f, $\{x,y\}$]* gives a list of vectors which represent monomials and correspond the monomials to the respective coefficient. If $f = x^2y + 6x^3y + 92y + 8x^4$ then the output would be $\{\{2, 1\} - > 1, \{3, 1\} - > 6, \{0, 1\} - > 92, \{4, 0\} - > 8\}$.

*Part[A,1]* is similar to the command *Take* but works in a more general fashion. For example if $A = \{\{2, 1\} - > 1, \{3, 1\} - > 6, \{0, 1\} - > 92, \{4, 0\} - > 8\}$ then the output will be $\{2, 1\} - > 1$.

*GroeberBasis[F,$\{x,y,z\}$]* returns a reduced Groebner basis of $F$ using lexicographic ordering with $x > y > z$.

# Bibliography

[1] David S. Dummit and Richard M. Foote. *Abstract Algebra.* John Wiley and Sons Inc., Hoboken, NJ, third edition, 2004.

[2] David Cox, John Little, and Donal O'Shea. *Ideal, Varieties, and Algorithms.* Springer, New York, NY, third edition, 2007.

[3] David Cox, John Little, and Donal O'Shea. *Using Algebraic Geometry.* Springer, New York, NY, second edition, 2005.

[4] Peter J. Olver, *Classical Invariant Theory.* Cambridge University Press, Cambridge, UK, 1999.

[5] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry* Springer, New York, NY, 1995.

[6] William Fulton. *Young Tableaux.* Cambridge University Press, Cambridge, UK, 1997.

[7] Bernd Sturmfels. *Algorithms in Invariant Theory.* SpringerWien, New York, NY, second edition, 2008.

[8] Theodore W. Gamelin. *Complex Analysis.* Springer. New York, NY, 2001.

[9] Bernd Sturmfels and Neil White. Gröbner Bases and Invariant Theory. IMA Preprint Series number 343, Minneapolis, Minnesota, 1987.

[10] David Eisendbud and Bernd Sturmfels. Finding sparse systems of parameters. *Journal of Pure and Applied Algebra,* **94**:143-157, 1994.

[11] Stefan Stiedel. Gröbner Bases of Symmetric Ideals. arxiv.org. 2012

[12] Ryan Shifler. Universal Groebner bases of Circulant Polynomial Systems. *Proceedings of the National Conference of Undergraduate Research.* Asheville, NC. 2011.

[13] William W. Adams and Philippe Loustaunau *An introduction to Groebner bases.* University Prees. Hyderabad, India. 1994.