APPLYING GRÖBNER BASES TO"SOLVE" SOLVABLE POLYNOMIALS

SAMUEL MOY

Abstract. This paper is designed to introduce the reader to Gröbner bases as well as demonstrate how they may applied in developing algorithms that"solve" solvable polynomials. Specifically, I will use these bases along with some basic invariant theory in order to derive a general formula for the roots of a cubic polynomial with rational coefficients and describe how to compute the exact roots of a quintic polynomial with rational coefficients with cyclic Galois group $\mathbb{Z}_5$.

Contents

1. Introduction

It has been nearly two centuries since the independent discoveries of two brilliant mathematicians, Niels Abel and variste Galois, put an end to the search for "general formulas" for polynomial roots by proving that no such "general formula" may exist for univariate polynomials of degree 5 or for polynomials of degree ≥ 5, respectively. Since then, the subject has received very little attention, despite the obvious fact that these proofs hardly lend themselves to finding the explicit roots of "solvable" polynomials (i.e. polynomials whose roots are expressible in terms of field elements, field operations, and radicals). In the hope of concretely soldifying the subject of solving univariate polynomials (with coefficients in $\mathbb{C}$), this paper presents an algorithm–using Gröbner bases and invariant theory–that allows one to actually "solve" the solvable polynomial. Before describing this computationally intensive algorithm, however, I first provide the audience with some background regarding Gröbner bases: monomial orders on polynomial rings, the computation of a Gröbner basis given a monomial order, and the properties and uses of Gröbner bases. Following this, I put forth some basic results and theorems in invariant

Date: 31 August 2010.

theory that are essential to understanding the algorithm. Last, I describe the algorithm in its generality before demonstrating how it is performed in a couple of particular cases.

2. Gröbner Bases

The key to the algorithm that will be presented is this paper is a special presentation of a set of generators for an ideal, known as a Gröbner basis. But before defining a Gröbner basis or describing the algorithm by which one may be obtained, we must first begin by defining a *monomial order*.

Definition 2.1. Let R be a ring. Consider the polynomial ring $R[x_1, x_2, \ldots, x_n] = R[\mathbf{x}]$. A *monomial order* $\succ$ is a total order on the set of all monic polynomials in $R[\mathbf{x}]$ that satisfies the following two properties:

i If u and v are monomials in $R[\mathbf{x}]$ such that $u \succ v$, then for any monomial w in $R[\mathbf{x}]$, we have $uw \succ vw$.

ii Every nonempty set of monomials in $R[\mathbf{x}]$ has a smallest element under $\succ$. Otherwise stated, $\succ$ is a well-ordering.

Example 2.2. Consider the polynomial ring $\mathbb{Z}[x]$. The only monomial order on $\mathbb{Z}[x]$ is given by $x^0 \prec x^1 \prec x^2 \prec x^3 \prec \cdots$ (It isn't difficult to arrive at a contradiction if we assume we have any other order, i.e. one in which $x^i \succ x^j$ for some $i < j$). If we consider some polynomial, $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then we know from elementary algebra that the leading term of this polynomial is simply the term with the highest power of x, in this case, $a_n x^n$.

Similarly, given any polynomial $q(\mathbf{x})$ in $R[\mathbf{x}]$ and a monomial order $\succ$, we may define a general notion of the leading term.

Definition 2.3. Let R a ring and consider the polynomial ring $R[\mathbf{x}]$ with $\succ$ as its monomial order. Let $p(\mathbf{x})$ be a polynomial in $R[\mathbf{x}]$ and let J be the chain with the monomial terms of $p(\mathbf{x})$ as its ordered elements. The *leading term* of $p(\mathbf{x})$, denoted by $LT(p)$, is the maximal element of the chain J.

Example 2.4. To bring a bit more clarity to the notion of the monomial order, let us consider some of the common monomial orders. Take R a ring and consider the polynomial ring $R[x_1, x_2, \ldots, x_n]$. Lexicographical order, abbreviated by "lex" and denoted by $x_1 \succ x_2 \succ \cdots \succ x_n$ (up to re-indexing the variables) works as follows:

The monomial with the highest power of x_1 is the largest, and, in the event of a "tie", the monomial of those in a "tie" with the highest power of x_2 is the largest. If a tie still remains after considering the x_i variable, the monomial of those in a "tie" with the highest power x_{i+1} is taken to be the largest.

While this purely lex order will be used for all of the calculations presented here, there are a number of other commonly used orders, including graded lexicographical order, and graded reverse lexicographical order. *Graded* means that the total degree of the monomial (the sum of the powers of all the x_i's) is considered before the lexicographical order is enforced. *Reverse* means that, given two polynomials of the same total degree, the polynomial with the *smaller* power of x_n is the larger monomial. In the event of a tie, the x_{n-1} variable is considered, and so on and so forth. It is of note that, empirically, it has been shown that most computations are done more quickly when grevlex (graded reverse lex) is the monomial order assigned to the polynomial ring.

With these definitions at our disposal, we now define a Gröbner basis as follows.

Definition 2.5. Let $R[\mathbf{x}]$ be a polynomial ring. Let $I = \langle f_1, f_2, \ldots, f_n \rangle$ be some finitely generated ideal. Define $LT(I) = \langle \{f | f \in I\} \rangle$, the ideal generated by the leading terms of the elements of I. A set of polynomials, $G = \{g_1, \ldots, g_k\}$ forms a *Gröbner basis* for I if $I = \langle G \rangle$ and $LT(I) = \langle LT(g_1), \ldots, LT(g_k) \rangle$.

Furthermore, our Gröbner basis is said to be *reduced* if $LT(g_i)$ does not divide any monomial occurring in g_j for all $i, j \in \{1, 2, \ldots, k\}, i \neq j$.

Remark 2.6. Although we refer to the previously defined object as a *basis*, it should be remembered that a Gröbner basis does not have all the properties that a vector space basis has. In particular, although a Gröbner basis generates its ideal I, an element in I is, in general, not generated *uniquely* by the elements of the Gröbner basis.

Notation 2.7. At this point, the usefulness of Gröbner bases may be far from clear. After all, a Gröbner basis just seems to be only one of a vast number of finite generating sets for an ideal. Nevertheless, they are distinguished by their exceptional applicability to performing computations. In fact, prior to their introduction to computational algebra in 1965 by Bruno Buchbereger, there were a number of seemingly simple calculations that could not be done. For example, computing a set of generators for the intersection of two ideals was an incredibly difficult calculation in the general case. This particular algorithm will be described later.

Theorem 2.8 (Hilbert's Monomial Ideal Theorem)**.** *Every monomial ideal $\mathcal{M}$ in $\mathbb{C}[x_1, x_2, \ldots, x_n]$ is finitely generated by monomials.*

Proofs of Hilbert's Monomial Ideal Theorem are generally included as part of the proof of the well-known Hilbert's Basis Theorem.

Theorem 2.9. *Let $\prec$ be a monomial order on $\mathbb{C}[\mathbf{x}]$. There is no infinite descending chain of monomials $m_1 \succ m_2 \succ m_3 \succ \ldots$.*

Proof. Let $\{m_1, m_2, m_3, \ldots\}$ be an infinite set of monomials in $\mathbb{C}[x_1, x_2, \ldots, x_n]$ with $\prec$ a monomial order. The ideal generated by this set is finitely generated by Hilbert's Monomial Ideal Theorem. Thus, there exists $j \in \mathbb{N}$ such that $m_j \in \langle m_1, m_2, \ldots, m_{j-1} \rangle$. Hence, m_i divides m_j for some $i < j$. Since $\prec$ is a monomial order, then $m_i \prec m_j$ for some $i < j$. Therefore, $\{m_1, m_2, m_3, \ldots\}$ cannot be an infinite descending chain of monomials under any monomial order. □

Remark 2.10. Given a subset of $R[x_1, x_2, \ldots, x_n]$, $\{f_1, f_2, \ldots, f_n\}$, how, then, can we compute a Gröbner basis?

The most straightforward algorithm simply requires multiple iterations of dividing f_j by f_i for some $i \neq j$ and then proceeding to replace f_j by the remainder obtained from that division. This procedure continues until the set of polynomials remains the same, regardless of the choice of i and j for division. Given a monomial order for our polynomial, then the previous theorem tells us that this algorithm will eventually terminate for any finite set of polynomials. If we do continue this process until it ends, then we ultimately obtain a reduced Gröbner basis. The ability to reduce our Gröbner bases to something of a minimal "size" is invaluable, as will soon be seen.

There are several theorems regarding Gröbner bases that have particular significance to our present endeavor.

Theorem 2.11. *Let I be an ideal and $\prec$ a monomial order on $\mathbb{C}[x_1, x_2, \ldots, x_n]$. The set of residue classes of monomials is a $\mathbb{C}$-vector space basis for the residue ring $\mathbb{C}[x_1, x_2, \ldots, x_n]/I$.*

Theorem 2.12 (Elimination)**.** *Consider the polynomial ring $\mathbb{C}[t, x_1, x_2, \ldots, x_n]$ with the purely lexicographical ordering $t \succ x_1 \succ x_2 \succ \cdots \succ x_n$. Let I be any ideal in this ring (It will necessarily be finitely generated by Hilbert's Basis Theorem). Let G be a Gröbner basis for this ideal. Then $I \cap \mathbb{C}[x_1, x_2, \ldots, x_n]$ is an ideal in $\mathbb{C}[x_1, x_2, \ldots, x_n]$, and $G \cap \mathbb{C}[x_1, x_2, \ldots, x_n]$ is a Gröbner basis for this ideal.*

Algorithm 2.13 (Intersection of Two Finitely Generated Ideals)**.** :
Let $A = \langle f_1, \ldots, f_r \rangle$ and $B = \langle g_1, \ldots, g_s \rangle$ be two finitely generated ideals from the Noetherian polynomial ring $R[x_1, \ldots, x_n]$.
Now, consider $I = \langle tf_1, tf_2, \ldots, tf_r, (1-t)g_1, \ldots, (1-t)g_s \rangle$, an ideal in $R[t, x_1, \ldots, x_n]$.
Claim: $A \cap B = I \cap R[x_1, \ldots, x_n]$. In other words, the first elimination ideal of I (elimination of the variable t) is equal to the intersection of the ideals A and B.

Proof. We must show both inclusions to proof the claim, and thereby validate the algorithm.
Suppose $f \in A \cap B$. Then $tf \in tA$ and $(1-t)f \in (1-t)B$, so

$$f = tf + (1-t)f \in tA + (1-t)B = I.$$

Suppose $f \in I \cap R[x_1, \ldots, x_n]$. Then

$$f = h_1 tf_1 + \ldots + h_r tf_r + h_{r+1}(1-t)g_1 + \ldots + h_{r+s}(1-t)g_s$$

where $h_1, \ldots, h_{r+s} \in R[t, x_1, \ldots, x_n]$. But the only way that we may then have $f \in R[x_1, \ldots, x_n]$ is if

$$f = h_{r+1}g_1 + \ldots + h_{r+s}g_s$$

where $h_{r+1}, \ldots, h_{r+s}$ are necessarily polynomials in $R[0, x_1, \ldots, x_n]$. This implies that $f \in B$. Finally, we are left with:

$$\begin{aligned} 0 &= (h_1 tf_1 + \ldots h_r tf_r) - (h_{r+1} tg_1 + \ldots h_{r+s} tg_s) \\ \Rightarrow h_1 tf_1 + \ldots h_r tf_r &= h_{r+1} tg_1 + \ldots h_{r+s} tg_s \\ \Rightarrow f &= h_1 f_1 + \ldots h_r f_r \end{aligned}$$

where $h_1, \ldots, h_r$ are polynomials in $R[0, x_1, \ldots, x_n]$. Hence, $f \in A$.
Therefore, $f \in A \cap B$, completing the proof. □

3. A Bit of Invariant Theory

Consider the polynomial ring $\mathbb{C}[\mathbf{x}]$, with $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Given some finite matrix group $\Gamma \subset GL(\mathbb{C}^n)$, invariant theory studies the polynomials that are invariant under the action of Γ, a set we denote by $\mathbb{C}[\mathbf{x}]^\Gamma$.

Definition 3.1 (Reynolds Operator)**.** In the course of studying the polynomials in $\mathbb{C}[\mathbf{x}]$ that are invariant under the action of some finite group Γ, it becomes very useful to have some way of generating an invariant, given any polynomial in $\mathbb{C}[\mathbf{x}]$. The map that we will refer to as the Reynolds Operator and denoted by "$*$" maps $\mathbb{C}[\mathbf{x}]$ surjectively onto $\mathbb{C}[\mathbf{x}]^\Gamma$. We define it as follows:

$$\begin{aligned} * : \mathbb{C}[\mathbf{x}] &\to \mathbb{C}[\mathbf{x}]^\Gamma \\ f &\mapsto f^* := \frac{1}{|\Gamma|} \sum_{\pi \in \Gamma} f \circ \pi \end{aligned}$$

Proposition 3.2. *The aforementioned Reynolds Operator "$*$" has the following properties.*

(a) "$*$" is a $\mathbb{C}$-linear map.
(b) "$*$" restricts to the identity map on $\mathbb{C}[\boldsymbol{x}]^\Gamma$.
(c) "$*$" is a $\mathbb{C}[\boldsymbol{x}]^\Gamma$-module homomorphism.

Theorem 3.3. *Hilbert's Finiteness Theorem The invariant ring $\mathbb{C}[\mathbf{x}]^\mathbf{\Gamma}$ of a finite matrix group $\Gamma \subset GL(\mathbb{C}^n)$ is finitely generated.*

Proof. Let $\mathcal{I}_\Gamma := \langle \mathbb{C}[\mathbf{x}]^\Gamma_+ \rangle$ be the ideal in $\mathbb{C}[\mathbf{x}]$ which is generated by all homogeneous invariants of positive degree. By Prop 2.2 (a), every invariant I is a $\mathbb{C}$-linear combination of symmetrized monomials $(x_1^{e_1}x_2^{e_2}\dots x_n^{e_n})$. These homogeneous invariants are the images of monomials under the Reynolds operator. Thus, $\mathcal{I}_\Gamma$ is generated by the polynomials $(x_1^{e_1}x_2^{e_2}\dots x_n^{e_n})^*$, where the vector $\mathbf{e} = (e_1, e_2, \dots, e_n)$ ranges over nonzero, nonnegative integer vectors.

By Hilbert's basis theorem, all ideals in the polynomial ring $\mathbb{C}[\mathbf{x}]$ are finitely generated. Hence, there exist finitely many homogeneous invariants $I_1, I_2, \dots, I_m$, such that $\mathcal{I}_\Gamma = \langle I_1, I_2, \dots, I_m \rangle$. Next, we prove that all homogeneous invariants $I \in \mathbb{C}[\mathbf{x}]^\Gamma$ can be written as polynomial functions in these I_j's.

Suppose not. Let I be a homogeneous element of minimal total degree in $\mathbb{C}[\mathbf{x}]^\Gamma \setminus \mathbb{C}[I_1, I_2, \dots, I_m]$. Since $I \in \mathcal{I}_\Gamma$, we have $I = \sum_{j=1}^{s} f_j I_j$ for some homogeneous polynomials $f_j \in \mathbb{C}[\mathbf{x}]$ of degree less than $\deg(I)$. Applying the Reynolds operator on both sides of this equation, we get:

$$I = I^* = (\sum_{j=1}^{s} f_j I_j)^* = \sum_{j=1}^{s} f_j^* I_j$$

by Proposition 2.2. The new f_j^*'s are homogeneous invariants whose degrees are all less than $\deg(I)$. By our assumption of minimality of the degree of I, we must have $f_j^* \in \mathbb{C}[I_1, I_2, \dots, I_m]$ for $j \in \{1, \dots, s\}$. But we then have $I \in \mathbb{C}[I_1, I_2, \dots, I_m]$, a contradiction! Thus, we have proved the theorem. □

As of now, however, we know only that our invariant ring is finitely generated. What we need to know for our later calculations is how to compute a set of fundamental invariants, given our polynomial ring, $\mathbb{C}[\mathbf{x}]$, and a finite group Γ acting on the ring.

Definition 3.4. A set of invariants $\{I_1, I_2, \dots, I_m\} \subset \mathbb{C}[\mathbf{x}]^\Gamma$ is said to be a *set of fundamental invariants* if it generates $\mathbb{C}[\mathbf{x}]^\Gamma$ as an algebra (over $\mathbb{C}$).

Proposition 3.5. *Let $I = \{I_1, I_2, \dots, I_m\}$ be a set of invariants in $\mathbb{C}[\boldsymbol{x}]^\Gamma$, such that there are no algebraic relations between the I_j's. Then we have the following equivalence:*

$$\mathbb{C}[I_1, I_2, \dots, I_n] = \mathbb{C}[\boldsymbol{x}]^\Gamma$$

Any set of invariants that satisfies the above equivalence is said to be complete. In particular, any set of fundamental invariants, assuming that it has been reduced so that there are no algebraic relations between the I_j's, is complete.

The following algorithm will allow us to compute a finite set of fundamental invariants for $\mathbb{C}[\mathbf{x}]^\Gamma$, assuming that our (finite) group is cyclic.

Algorithm 3.6 (Computing fundamental invariants for a finite abelian group)**.** Let $\Gamma \subset GL(\mathbb{C}^n)$ be a finite cyclic group. Let Ω be a generating matrix for Γ.
Step 1: Using eigenvectors for Ω, find a matrix T_Γ which diagonalizes Ω.
Step 2: Introduce new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ by setting $\mathbf{y} = T_\Gamma \mathbf{x}$.
Step 3: Write $D = T_\Gamma \Omega T_\Gamma^{-1} = \text{diag}(\omega_1, \omega_2, \ldots, \omega_n)$. In the case of our cyclic group, the ω_j's will all be n^{th} roots of unity. Let d_i denote the smallest positive integer such that $\omega_i = \zeta^{d_i}$, where ζ is some primitive n^{th} root of unity (and, hence, a generator). Let g denote the order of the matrix D (i.e. g is the smallest positive integer such that $D^g = I_n$). Since Γ is cylic, we will have $g = |\Gamma|$.
Step 4: Consider the linear homogenous congruence:

$$d_1\mu_1 + d_2\mu_2 + \cdots + d_n\mu_n \equiv 0 \mod g$$

Compute a finite generating set $\mathcal{H}$ for the solution monoid of the system. Then:

$$\mathbb{C}[\mathbf{x}]^\Gamma = \mathbb{C}[\mathbf{y}]^{T_\Gamma \Omega T_\Gamma^{-1}} = \mathbb{C}[y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n} \mid \mu = (\mu_1, \mu_2, \ldots, \mu_n) \in \mathcal{H}].$$

Explanation 3.7 (Algorithm 3.6). Why exactly does this algorithm work?

To begin, an invariant under the action of Γ is the same as an eigenfunction with eigenvalue 1. By changing basis from the x_i's to the y_i's (which are degree 1 eigenfunctions), we have a basis of eigenfunctions given by monomials in the y_i's. In order to determine a set of generators for the ring of invariances, $\mathbb{C}[\mathbf{x}]^\Gamma$, it suffices to find a set of monic monomial generators for this ring, each of the form $y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n}$. For an arbitrary monomial of this form to be invariant under the action of our group Γ, it must be invariant under the action of the generating matrix Ω and, hence, also invariant under the action of our diagonal matrix D. Furthermore, given a monomial $\mathscr{M}_\mathbf{y} = y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n}$, $\mathscr{M}_\mathbf{y}$ is invariant under the action of D if and only if $D \cdot \mathscr{M}_\mathbf{y} = \mathscr{M}_\mathbf{y}$ if and only if the eigenvalue of the monomial $\mathscr{M}_\mathbf{y}$ is 1.

Now, for our diagonal matrix D, the eigenvalue of y_i is ω_i for $i \in \{1, 2, \ldots, n\}$. Hence, because Γ is cyclic and, in particular, abelian, the eigenvalue (under the action of D) of an arbitrary monomial of the form $y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n}$ is:

$$\lambda_{y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n}} = \lambda_{y_1^{\mu_1}} \lambda_{y_2^{\mu_2}} \ldots \lambda_{y_n^{\mu_n}} = (\lambda_{y_1})^{\mu_1} (\lambda_{y_2})^{\mu_2} \ldots (\lambda_{y_n})^{\mu_n} = \omega_1^{\mu_1} \omega_2^{\mu_2} \ldots \omega_n^{\mu_n}$$

Since each of the ω_i's is a root of unity, we may express each in terms of some chosen primitive n^{th} root of unity, ζ. Hence, $\omega_i = \zeta^{d_i}$ for $i \in \{1, 2, \ldots, n\}$. The monomials that are invariant under our group will those such that:

$$\begin{aligned}
&\lambda_{y_1^{\mu_1} y_2^{\mu_2} \ldots y_n^{\mu_n}} = 1 \\
\Leftrightarrow &\lambda_{y_1^{\mu_1}} \lambda_{y_2^{\mu_2}} \ldots \lambda_{y_n^{\mu_n}} = 1 \\
\Leftrightarrow &\omega_1^{\mu_1} \omega_2^{\mu_2} \ldots \omega_n^{\mu_n} = 1 \\
\Leftrightarrow &(\zeta^{d_1})^{\mu_1} (\zeta^{d_2})^{\mu_2} \ldots (\zeta^{d_n})^{\mu_n} = 1 \\
\Leftrightarrow &\zeta^{d_1 \cdot \mu_1} \zeta^{d_2 \cdot \mu_2} \ldots \zeta^{d_n \cdot \mu_n} = 1 \\
\Leftrightarrow &\zeta^{d_1 \cdot \mu_1 + d_2 \cdot \mu_2 + \ldots d_n \cdot \mu_n} = 1 \\
\Leftrightarrow &d_1\mu_1 + d_2\mu_2 + \ldots + d_n\mu_n \equiv 0 \mod g \\
&\text{(where g is the order of the cyclic group } \Gamma)
\end{aligned}$$

Now, take $\mathcal{H}$ to be a generating set for the solution monoid defined by:

$$d_1\mu_1 + d_2\mu_2 + \ldots + d_n\mu_n \equiv 0 \mod g.$$

Then the set $\{y_1^{\mu_1}y_2^{\mu_2}\dots y_n^{\mu_n} \mid \mu = (\mu_1, \mu_2, \dots, \mu_n) \in \mathcal{H}\}$ will generate all monomials invariant under the action of the group Γ. Therefore, we will have found generators for the ring of invariances:

$$\mathbb{C}[\mathbf{x}]^\Gamma = \mathbb{C}[\mathbf{y}]^{T_\Gamma \Omega T_\Gamma^{-1}} = \mathbb{C}[y_1^{\mu_1}y_2^{\mu_2}\dots y_n^{\mu_n} \mid \mu = (\mu_1, \mu_2, \dots, \mu_n) \in \mathcal{H}].$$

Notation 3.8. In general, the computation of fundamental invariants is significantly more difficult than in the cyclic case. Nevertheless, algorithms (relying on the Reynolds operator) do exist for computing the fundamental invariants for the invariant ring $\mathbb{C}[\mathbf{x}]^\Gamma$ for any finite matrix group $\Gamma \subset GL(\mathbb{C}^n)$.

4. Solving "Solvable" Polynomials

Classical Galois theory tells us that a polynomial with rational coefficients can be solved if and only if its corresponding Galois group is solvable. That said, determining how to find the exact roots of a polynomial with solvable Galois group is a task that is rarely considered. This section will describe how to arrive at a general formula for finding the roots of polynomials with particular degree and Galois group. Afterward, the process will be applied to polynomials of degree 3, and the result of the process will be described for polynomials of degree 5 with cyclic Galois group.

First, a quick note on the elementary symmetric polynomials.

Notation 4.1 (Elementary Symmetric Polynomials)**.** Suppose we consider the polynomial ring $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, x_2, \dots, x_n]$. Then there are n elementary symmetric polynomials, and for $k \in \{1, 2, \dots, n\}$, $\sigma_k(\mathbf{x})$ is defined as follows:

$$\sigma_k(\mathbf{x}) = \sum_{1 \le j_1 < j_2 < \dots < j_k \le n} x_{j_1}x_{j_2}\dots x_{j_k}$$

The elementary symmetric polynomials have a number of desirable properties. The most significant is the fact that the elementary symmetric polynomials generate all symmetric polynomials in a polynomial ring. In other words, given $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, x_2, \dots, x_n]$ and the group S_n acting in the natural way on the variables $x_1, x_2, \dots, x_n$, then $\mathbb{C}[\mathbf{x}]^{S_n} = \mathbb{C}[\sigma_1(\mathbf{x}), \sigma_2(\mathbf{x}), \dots, \sigma_n(\mathbf{x})]$.

Now, to begin, I will describe the general idea of computing the roots of a polynomial in $\mathbb{Q}[z]$ with solvable Galois group Γ. Let us assume we have a polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z_1 + a_0$ with roots $x_1, x_2, \dots, x_n$. Then, equivalently, we have $p(z) = (z - x_1)(z - x_2)\dots(z - x_n)$. By expanding and setting the two expressions equal, we obtain the following equations:

$$\begin{aligned}
\sigma_1(\mathbf{x}) + a_{n-1} &= 0 \\
\sigma_2(\mathbf{x}) - a_{n-2} &= 0 \\
&\vdots \\
\sigma_j(\mathbf{x}) - (-1)^j a_{n-j} &= 0 \\
&\vdots \\
\sigma_n(\mathbf{x}) - (-1)^n a_0 &= 0
\end{aligned}$$

If we consider the ideal $I = \langle \sigma_1(\mathbf{x}) + a_{n-1}, \sigma_2(\mathbf{x}) - a_{n-2}, \dots, \sigma_n(\mathbf{x}) - (-1)^n a_0 \rangle$, then the above equations are effectively "encoded" into the affine variety $\mathcal{V}(I)$.

However, in general, these relations alone are not sufficient for us to determine the roots of the polynomial. We must now consider the Galois group Γ. Since Γ is a finite solvable group, then we may find a composition series

$$\Gamma = \Gamma_1 \triangleright \Gamma_2 \triangleright \ldots \triangleright \Gamma_{k-1} \triangleright \Gamma_k = \{e\}$$

satisfying the property that Γ_i/Γ_{i-1} is cyclic of prime order p_i for $i \in \{1, 2, \ldots, k\}$.

The following steps "work backwards" in order to encode the action (by the Galois group Γ) on the roots into our affine variety.

Step 1: For the last non-trivial group in our series, Γ_{k-1}, compute (generators for) the invariant ring $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-1}}$.

Step 2: Compute the invariant ring $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-2}}$ by considering the action of $\Gamma_{k-2}/\Gamma_{k-1} = \mathbb{Z}_{p_{k-2}}$ on $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-1}}$. This procedure is justified because $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-1}}$ is a free module of rank p_{k-2} over $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-2}}$. In particular, one need only consider the group action of $\mathbb{Z}_{p_{k-2}}$ on the generators of $\mathbb{C}[\mathbf{x}]^{\Gamma_{k-1}}$. Explicitly, the action of $\mathbb{Z}_{p_{k-2}}$ on a generator $\xi_j \in \mathbb{C}[\mathbf{x}]^{\Gamma_{k-1}}$ is given by:

$$\xi_j \mapsto \zeta_{k-2} \cdot \xi_j$$

for each of the ξ_j's, where ζ_{k-2} is a primitive p_{k-2}^{th} root of unity.

Step 3: Continue determining a set of generators for each $\mathbb{C}[\mathbf{x}]^{\Gamma_i}$ in this manner by similarly considering the group action of Γ_i/Γ_{i+1} on the generators of the invariant ring $\mathbb{C}[\mathbf{x}]^{\Gamma_{i+1}}$. This will provide us with a set of generators for $\mathbb{C}[\mathbf{x}]^{\Gamma}$.

Step 4: Having determined a set of generators for $\mathbb{C}[\mathbf{x}]^{\Gamma}$, restrict this invariant ring to the affine variety $\mathcal{V}(I)$. Computing a Gröbner basis once this restriction is made will output a basis with the generators for $\mathbb{C}[\mathbf{x}]^{\Gamma}$ in terms of the a_i's. If explicit a_i's are given, then the exact values of the generators for $\mathbb{C}[\mathbf{x}]^{\Gamma}$ may be determined (using only the "general formulas" for the proper subgroups of Γ). With these, we may explicitly determine the values of the generators of $\mathbb{C}[\mathbf{x}]^{\Gamma_2}$, which may be used to determine the values of the generators of $\mathbb{C}[\mathbf{x}]^{\Gamma_3}$, etc. Continuing in this procedure, we will determine the exact values of the generators for each of the $\mathbb{C}[\mathbf{x}]^{\Gamma_i}$'s.

Step 5: Finally, we explicitly determine the roots $x_1, x_2, \ldots, x_n$ in the final step going from $\Gamma_{k-1}/\Gamma_k = \Gamma_{k-1}$ to $\Gamma_k = \{e\}$.

In order to demonstrate this process in action, I will perform the necessary computations to determine the general formula for cubic equations. Additionally, though the calculations for quintic polynomials with cyclic Galois group $\mathbb{Z}_5$ proved too computationally taxing, I will describe the conjectured result.

Example 4.2 (Derivation of Cardano's Formula for Cubics)**.** Let us take an arbitray monic cubic polynomial in $\mathbb{C}[z]$:

$$p(z) = z^3 + a_2 z^2 + a_1 z + a_0 = (z - x_1)(z - x_2)(z - x_3)$$

In order to arrive at a general formula, we must express each of the x_i's in terms of only the a_j's. The symmetric relations for this polynomial are the following: $x_1+x_2+x_3 = -a_2$, $x_1x_2+x_1x_3+x_2x_3 = a_1$, and $x_1x_2x_3 = -a_0$. These relations may be encoded in the ideal: $\mathcal{I} = \langle x_1+x_2+x_3+a_2, x_1x_2+x_1x_3+x_2x_3-a_1, x_1x_2x_3+a_0\rangle$. Obviously, the solution $\mathbf{x} = (x_1, x_2, x_3)$ is contained in the affine variety $\mathcal{V}(I)$; however, the Galois group must be considered in order to determine the solution exactly. For the sake of generality, I assume the Galois group of $p(z)$ to be S_3.

The composition series for S_3 is simply $S_3 \triangleright A_3 \triangleright \{e\}$. So, we must first determine

generators for $\mathbb{C}[\mathbf{x}]^{A_3}$. Following Algorithm 2.6 (since A_3 is cyclic), we introduce the the variables $\mathbf{y} = (y_0, y_1, y_2)$ and obtain the following relations:

$$y_0 = x_1 + x_2 + x_3, \quad y_1 = x_1 + \zeta^2 x_2 + \zeta x_3, \quad y_2 = x_1 + \zeta x_2 + \zeta^2 + x_3$$

where ζ is a primitive 3^{rd} root of unity. After being "inverted" so that each of x_1, x_2, and x_3 are in terms of the y_i's, these relations allow for the Gröbner basis:

$$\mathcal{G}_0 = \{3x_1 - y_0 - y_1 - y_2, 3x_2 - y_0 - \zeta y_1 - \zeta^2 y_2, 3x_3 - y_0 - \zeta^2 y_1 - \zeta^2 y_2\}.$$

Next, following through with Algorithm 2.6 gives us four generators, $u_0 = y_0$, $u_1 2 = y_1 y_2$, $u_1 3 = y_1^3$, and $u_2 3 = y_2^3$, for the invariant ring $\mathbb{C}[\mathbf{x}]^{A_3}$. Taking a Gröbner basis of the set $\{u_0 - y_0, u_{12} - y_1 y_2, u_{13} - y_1^3, u_{23} - y_2^3\}$ gives:

$$\mathcal{G}_1 = \{u_0 - y_0, u_{12} - y_1 y_2, u_{13} - y_1^3, u_{23} - y_2^3, y_1^2 u_{12} - y_2 u_{13},$$
$$y_1 u_{12}^2 - y_2^2 u_{13}, y_1 u_{23} - y_2^2 u_{12}, u_{12}^3 - u_{13} u_{23}\}.$$

After this, we must consider the action of $S_3/A_3 \cong \mathbb{Z}_2$ on the roots of the polynomial and its effect upon the generators for $\mathbb{C}[\mathbf{x}]^{A_3}$. It is not difficult to see that any of the non-trivial actions of $\mathbb{Z}_2$ on x_1, x_2, and x_3 (i.e. x_1 and x_2 are "switched", etc.) only serves to permute y_1 with y_2, or, in terms of the u_i's, permute u_{13} with u_{23}. A generating set for $\mathbb{C}[\mathbf{x}]^{(S_3/A_3)\cdot A_3} = \mathbb{C}[\mathbf{x}]^{S_3}$ is, therefore, given by four generators, $v_0 = u_0$, $v_{12} = u_{12}$, $v_{13} = u_{13} + u_{23}$, and $v_{23} = u_{13} u_{23}$, which are encoded in the Gröbner basis:

$$\mathcal{G}_2 = \{v_{23} - u_{23} v_{13} + u_{23}^2, v_{13} - u_{13} - u_{23}, v_{12} - u_{12}, v_0 - u_0\}.$$

Before we compute a Gröbner basis for the relative orbit variety, $\mathcal{V}(I)/S_3$, we must transform the generators for the ideal $\mathcal{I}$ from being in terms of x_i's to being in terms of y_j's (in order to avoid messy computation with roots of unity). Hence, using $\mathcal{G}_0$, we determine:

$$\mathcal{I} = \langle y_0 + a_2, \frac{1}{3} y_0^2 - \frac{1}{3} y_1 y_2 - a_1, \frac{1}{27} y_0^3 + \frac{1}{27} y_1^3 + \frac{1}{27} y_2^3 - \frac{1}{9} y_0 y_1 y_2 + a_0 \rangle.$$

Finally, we compute the relative orbit variety, $\mathcal{V}(I)/S_3$, by computing a Gröbner basis for $\mathcal{I} \cup \mathcal{G}_1 \cup \mathcal{G}_2$ and then eliminate the variables $y_0, y_1, y_2, u_0, u_{12}, u_{13}, u_{23}$, resulting in the following output:

$$\mathcal{G} = \{v_{23} - a_2^6 + 27 a_1^3, v_{13}, v_{12} - a_2^2 + 3 a_1, v_0 + a_2\}.$$

Thus, we see that $\mathcal{G}$ allows us to determine each of the v's solely in terms of the a's. The v's, in turn, allow us to solve for the u's, which allow us to solve for the y's, which allow us to finally solve for x_1, x_2, x_3. Best of all, we need only be able to extract cube roots, square roots, or to solve polynomial equations of "lesser" degree (i.e. a quadratic). If all of these substitutions (which can, theoretically, be done by hand) were made, the end result would be Cardano's formula for cubic polynomials.

Example 4.3 (Quintic Polynomials with Cyclic Galois Group $\mathbb{Z}_5$)**.** Given an arbitrary fifth degree univariate polynomial, we may use the same steps to determine the relative orbit variety, $\mathcal{V}(I)/\mathbb{Z}_5$. However, lacking the Gröbner basis for the relative orbit variety, we may only put forth a conjecture regarding this missing output

Since $\mathbb{Z}_5$ is cyclic, we must only determine generators for the invariant ring once. Nevertheless, since $\mathbb{Z}_5 \subset S_5$ implies that $\mathbb{C}[\mathbf{x}]^{S_5} \subset \mathbb{C}[\mathbf{x}]^{\mathbb{Z}_5}$, then the computed Gröbner basis for the relative orbit variety should not allow us to determine a "general formula" for quintics with cyclic Galois group, since it is well known

that no "general formula" for quintics can exist. Rather, it seems far more likely that the compued Gröbner basis will have polynomials that are only solvable given "correct" a_j's from a quintic polynomial with cyclic Galois group. In essence, once the coefficients from a polynomial with cyclic Galois group are substituted, the polynomials in the Gröbner basis will "simplify" to forms such that they may be solved only by extracting fifth roots and solving (solvable) polynomials of "lesser" degree.

5. Conclusion

Although Gröbner basis computations are very "costly" from a computer science standpoint, their application to solving for explicit polynomial roots brings life to a subject long thought to be quite dead. Despite the abstract nature of Galois' work, the algorithm presented in this project finally brings forth new, concrete meaning to the term "solvable."

6. Acknowledgments

I would like to express my sincere thanks to my graduate mentors, Aaron Marcus and Emily Norton. I am particularly indebted to Aaron for his extreme patience with my frequent bouts of procrastination and for giving me the opportunity to experience his overall, general awesomeness. Additionally, I would like to thank Peter May for, yet again, spearheading a program so very fruitful and rewarding as the UChicago mathematics REU.

References

[1] Sturmfels, Bernd. "Algorithms in Invariant Theory." New York: SpringerWien, 1993.
[2] Cox, David, John B. Little, and Don O'Shea. "Ideals, Varieties, and Algorithms." New York: Springer, 2007.
[3] Becker, Thomas, Weispfenning, Volker. "Gröbner Bases: A Computational Approach to Commutative Algebra." Berlin: Springer, 1993.