AUTOMATIC GEOMETRY THEOREM PROVING

AMOL RAMA AND NILAY MISHRA

July 13, 2020

ABSTRACT. A typical Olympiad geometry problem involves complex configurations with well-defined points but nearly impossible to guess synthetic observations. As a result, many of the contestants use some form of algebra to assist them in their endeavors. In a similar way, we can use Wu's Method, which is based on Ritt's Principle and pseudodivision, to solve many of these challenging problems using a computer.

1. INTRODUCTION

The goal of most olympiad questions (especially in the USAMO or IMO) is to create insanely challenging questions, that require nothing more than basic high school mathematics to solve. In particular, these problems always have a solution that refrain from using advanced methods such as calculus, abstract algebra, or linear algebra. On the flip side, certain questions are notorious for being easier with higher methods (e.x. 2008 USAMO/6), and popular techniques for solving inequalities (e.x. Jensen's inequality, Lagrange multipliers, tangent line trick, etc.) are based upon calculus.

These questions are nevertheless quite challenging at the upper end of the spectrum, requiring immense observation, creativity, and sometimes computational fortitude to solve. A key example of this is Olympiad geometry questions. While these never require a result beyond high school mathematics to solve (e.x. algebraic geometry), they can be immensely difficult to the untrained eye and can bury problem-solvers with complicated diagrams and configurations.

While the general *synthetic* way to solve these problems is by using only basic facts of classical Euclidean geometry, some contestants opt to solve geometry problems in a computational manner. Here are some possible computational methods of solving almost any geometry problem of Olympiad caliber:

- Trigonometry
- Cartesian coordinates
- Complex numbers
- Barycentric (also called areal or trillinear) coordinates

Some good places to refer to for learning these methods are Euclidean Geometry in Mathematical Olympiads, Barycentric Coordinates in Olympiad Geometry, and Lemmas in Olympiad Geometry.

IMO problems, and Olympiad problems in general, have often been characterized as "not straightforward" and requiring an immense amount of creativity and practice to come up with synthetic solutions. Consider the IMO Grand Challenge (Link this to page), whose goal is to advance artificial intelligence to the point where a computer can solve arbitrary Olympiad questions in the same way as the top IMO contestants can, and earn enough points to receive a "gold medal."

The idea behind this paper is to present the reader with an algorithm, that can be implemented into a computer with a variety of programming languages, that can tackle almost any IMO geometry problem in 2D space. This is motivated by the simple question, *We don't **need** advanced methods such as algebraic geometry, but what if we use them?*. The answer to this question is that with results from algebraic geometry, we can simulate a computational solution with the additional benefit that the computational fortitude of most computers far exceed that of a typical contestant.

What makes geometry special as compared to other subjects when solving Olympiad problems? The idea behind this is that any geometry problem can be solved given enough time and effort to expend, using a computational

method. Most Olympiad contestants know this intrinsically, but this idea of solving creative problems using brute force is not unique to geometry (see SOS method, inequalities [5]). In fact, the key idea is the following:

Theorem 1.1 (Tarski)
Euclidean geometry is decidable i.e. there exists an algorithm to determine whether an arbitrary statement is true or not.

Using Wu's method, the computer outputs a result that may be adjusted to be human-readable (see this), but is often just a verification of the result (i.e. checking to see if the statement is true).

What is Wu's algorithm? That's the question this paper is trying to answer, but let's start off with a motivating example.

Consider the following problem:

Motivating Example 1
Let ABC be an arbitrary triangle, and let A', B', and C' lie in the exterior of the triangle, such that $A'BC$, $B'AC$, and $C'AB$ are all equilateral. Prove that AA', BB', and CC' are all concurrent.

An experienced contestant will instantly deduce that the point of concurrence is the Fermat point F, which is the point that minimizes $AF + BF + CF$ in a triangle. The result can be shown synthetically using Jacobi's theorem, or computationally using trigonometry, complex numbers, and barycentric coordinates. Let's see how we can use Wu's algorithm to automate proving this result.

Let's try to prove this result using Cartesian coordinates. Let $A = (0, 0)$, $B = (u_1, 0)$, and $C = (u_2, u_3)$. Let $C' = (x_1, x_2)$, $B' = (x_3, x_4)$, $A' = (x_5, x_6)$, and $S = (u_4, u_5)$. We have the following hypotheses:

- C' is defined as the point outside of the triangle for which ABC' is equilateral, so we have:

$$x_2 < 0$$

$$u_1^2 = x_1^2 + x_2^2$$

$$u_1^2 = (u_1 - x_1)^2 + x_2^2$$

- Similarly, B' is defined as the point outside the triangle for which $B'AC$ is equilateral, so we have:

$$x_4 > \frac{x_3 u_3}{u_2}$$

$$u_2^2 + u_3^2 = x_3^2 + x_4^2$$

$$u_2^2 + u_3^2 = (x_3 - u_2)^2 + (x_4 - u_3)^2$$

- A' is defined as the point outside the triangle for which $A'BC$ is equilateral, so we have:

$$x_6 > \frac{u_2 - u_1}{u_3} x_5 + \frac{u_1^2 - u_1 u_2}{u_3}$$

$$(u_1 - u_2)^2 + u_3^2 = (x_5 - u_1)^2 + x_6^2$$

$$(u_1 - u_2)^2 + u_3^2 = (x_5 - u_2)^2 + (x_6 - u_3)^2$$

- C, S, and C' are collinear, so:

$$u_4 x_2 + u_3 x_1 - u_5 x_1 - u_2 x_2 - u_3 u_4 + u_2 u_5 = 0$$

- B, S, and B' are collinear, so:

$$u_4 x_4 + u_2 x_3 - u_5 x_3 - u_1 x_4 - u_2 u_4 + u_1 u_5 = 0$$

Note that we have a total of 11 hypotheses. Number these as:

$$f_1 : x_2 < 0$$
$$f_2 : u_1^2 = x_1^2 + x_2^2$$
$$f_3 : u_1^2 = (u_1 - x_1)^2 + x_2^2$$
$$f_4 : x_4 > \frac{x_3 u_3}{u_2}$$
$$f_5 : u_2^2 + u_3^2 = x_3^2 + x_4^2$$
$$f_6 : u_2^2 + u_3^2 = (x_3 - u_2)^2 + (x_4 - u_3)^2$$
$$f_7 : u_4 x_2 + u_3 x_1 - u_5 x_1 - u_2 x_2 - u_3 u_4 + u_2 u_5 = 0$$
$$f_8 : u_4 x_4 + u_2 x_3 - u_5 x_3 - u_1 x_4 - u_2 u_4 + u_1 u_5 = 0$$
$$f_9 : x_6 > \frac{u_2 - u_1}{u_3} x_5 + \frac{u_1^2 - u_1 u_2}{u_3}$$
$$f_{10} : (u_1 - u_2)^2 + u_3^2 = (x_5 - u_1)^2 + x_6^2$$
$$f_{11} : (u_1 - u_2)^2 + u_3^2 = (x_5 - u_2)^2 + (x_6 - u_3)^2$$

We wish to prove that A, S, and A' are collinear, namely:

$$g : u_4 u_6 - x_5^2 = 0$$

In a more abstract sense, the first step of Wu's algorithm is to convert a geometry theorem into a set of algebraic equations $f_1, f_2, \cdots, f_r$ (our example had $r = 11$), and a target conclusion equation g. The goal of the algorithm is to extrapolate the given equations and figure out whether or not the conclusion follows generally from the hypothesis or not. This will be made more clear later on.

What happens next? We convert our system into a *triangular form*. In particular, we can write our system in the form:

$$f_1 = f_1 (u_1, u_2, \cdots, u_{d-1}, u_d, x_1)$$
$$f_2 = f_2 (u_1, u_2, \cdots, u_{d-1}, u_d, x_1, x_2)$$
$$f_3 = f_3 (u_1, u_2, \cdots, u_{d-1}, u_d, x_1, x_2, x_3)$$
$$\cdots$$
$$f_{r-1} = f_{r-1} (u_1, u_2, \cdots, u_{d-1}, u_d, x_1, x_2, x_3, \cdots, x_{r-2}, x_{r-1})$$
$$f_r = f_r (u_1, u_2, \cdots, u_{d-1}, u_d, x_1, x_2, x_3, \cdots, x_{r-1}, x_r)$$

in such a way that the variety $V(f_1, f_2, \cdots, f_r)$ contains the irreducible components of the original variety defined by the hypotheses, through the algorithm. Afterwards, we will perform pseudodivision (see Section 2), successively on the neew hypotheses in triangular form and the conclusion g. This yields a final remainder r, which is 0 if and only if g "follows" from $f_1, f_2, \cdots, f_r$. We will make this more rigorous later on.

Lastly, we will examine the nondegenerate conditions that arise during triangulation. We can conclude that g follows from the hypotheses $f_1, \cdots, f_r$ given that the nondegenerate conditions hold. In particular, these conditions take the form $p \neq 0$ where p is a polynomial that arises naturally during our triangulation process.

To summarize, we have the following process, roughly speaking. We will make this more rigorous for the remainder of the paper.

(1) Translate the geometry theorem into a set of algebraic equations and a target conclusion equation.
(2) Write the hypotheses in triangular form.
(3) Perform successive pseudodivision, and compute the remainder that results between the hypotheses $f_1, f_2, \cdots, f_r$
(4) Conclude, if possible, using nondegenerate conditions.

This method allows to prove or disprove any geometry theorem, including those at an Olympiad caliber of difficulty. This is the central driving force behind this paper. In the following sections, we will examine each step of the process in a more detailed fashion.

In order to determine if a result follows, we need to manipulate the expressions given to reduce the unecessary fat. For example, I could tell you that $x = 1$ and $(x-2)^2 = 1$, but we don't actually need the second piece of information. That's the idea behind pseudodivision - we get rid of unnecessary blubber in this fashion.

Definition 1 (Pseudodivision) — For two polynomials $f, g \in k[x_1, \ldots, x_n, y]$ with $f = \sum_{i=1}^{n} a_i y^i$ and $g = \sum_{i=1}^{m} b_i y^i$, there is an equation

$$a_n^s g = qf + r$$

where $s \geq 0$, $q, r \in k[x_1, \ldots, x_n, y]$, and r has degree less than n. Furthermore, r lies in the ring (f, g) in $k[x_1, \ldots, x_n, y]$. r is also typically denoted as $\mathrm{pr}(f, g, y)$.

This is equivalent to polynomial division, and r is also known as the "remainder" although we will continue to use $\mathrm{pr}(f, g, y)$.

It would be nice if we were only given one constraint, as then we perform this division exactly once and then we are done. However, we're normally given multiple conditions, and thus we will have to instead do continuous pseudodivision.

Definition 2 (Continuous Pseudodivision) — Consider some polynomials $g \in k[u_1, \ldots, u_d, x_1, \ldots, x_r]$ and $f_k \in k[u_1, \ldots, u_d, x_1, \ldots, x_k]$, where the f_k are all in triangular form. We define the sequence $\{R_k\}_{k=0}^{r}$ such that

$$R_r = g \qquad R_k = \mathrm{pr}(R_{k+1}, f_{k+1}, x_{k+1})$$

For the sake of convenience, define $R_0 \equiv \mathrm{pr}(g, f_1, \ldots, f_r)$.

Let's look at a few results that involve pseudodivision that will come in handy later:

Proposition 2.1

Consider some polynomials $g \in k[u_1, \ldots, u_d, x_1, \ldots, x_r]$ and $f_k \in k[u_1, \ldots, u_d, x_1, \ldots, x_k]$, where the f_k are all in triangular form and let $R = \mathrm{pr}(g, f_1, \ldots, f_r)$. Furthermore, let d_j be the leading coefficient of f_j (as viewed in a polynomial in x_j). Then, we can say that there exist nonnegative integers $s_1, \ldots, s_r$ and polynomials $A_1, \ldots, A_r$ such that

$$g \prod_{k=1}^{r} d_k^{s_k} = R + \sum_{k=1}^{r} A_k f_k$$

and either $R \equiv 0$ or $\deg(R, x_k) < \deg(f_k, x_k)$ for all $1 \leq k \leq n$.

Proof. We shall establish this by induction on r.
Base Case. $r = 1$. In this case, the result stated follows from our initial definition of pseudodivision, which in turn followed from polynomial division.
Induction Hypothesis. Assume the statement is true for $r - 1$, where $r \geq 2$. Then, we shall show it is true for r as well.
Induction Step. Note that by the induction hypothesis on the first $r - 1$ of the f_k, we have

$$R_{r-1} \prod_{k=1}^{r-1} d_k^{s_k} = R + \sum_{k=1}^{r-1} A_k f_k$$

By the $r = 1$ case, we can write $d_r^{s_r} g = R_r + A_r f_r$, and by substitution, we reach our original statement (as $R_r = g$). For the second part, note that $\deg(R_r, x_r) < \deg(f_r, x_r)$ by the $r = 1$ case, and the rest of the inequalities follow from the induction hypothesis. □

Pseudodivision highly depended on polynomials in triangular form. Although it **is** possible for us to use pseudodivision recursively without triangular form, it is much more useful to use the triangular form. Thus, the main challenge would be to put our polynomials into triangular form, motivating the following question: What's the smallest extension of k we can put f in, for some polynomial f? To answer this, we have the following definition:

Definition 3 (Class) — The class of a polynomial f (denoted as $\text{class}(f)$) is the smallest integer c such that $f \in k[x_1, \ldots, x_c]$. In particular, if $f \in k$, $\text{class}(f) = 0$.

For sake of convenience, we denote $x_{\text{class}(f)}$ as the leading variable, or $\text{LV}(f)$, and the leading coefficient of f when viewed in $\text{LV}(f)$ as $\text{LC}(f)$ (called leading coefficient). Finally, let the degree of f as viewed in $\text{LV}(f)$ be $\text{LD}(f)$. To help us simplify, we use the following notation:

Definition 4 (Reduced Polynomials) — We call a polynomial g **reduced** with respect to f if $\deg(g, x_c) < \text{LD}(f)$, where $c = \text{class}(f)$.

One useful way of thinking about this is $g = \text{pr}(g, f, x_c)$. As a corollary to successive pseudodivision, we introduce reducedness to a set of polynomials:

Definition 5 (Reduced Polynomials, Extended) — We say that a polynomial $g \in k[u_1, \ldots, u_d, x_1, \ldots, x_r]$ is reduced to the trianglular polynomials $f_1, \ldots, f_r$ with $f_c \in k[u_1, \ldots, u_d, x_1, \ldots, x_c]$ if $\deg(g, x_i) < \deg(f_i, x_i)$ for all $1 \leq i \leq n$.

Now, continuing upon this, we shall talk about ascending chains:

Definition 6 (Ascending Chains) — Consider a sequence of polynomials $f_1, \ldots, f_r$ in $k[x_1, \ldots, x_n]$. We call it an ascending chain if $r = 1$ and $f_1 \neq 0$, or $r > 1$ and

$$0 < \text{class}(f_1) < \text{class}(f_2) < \cdots < \text{class}(f_r)$$

Furthermore, we must have for all $i < j$, f_j is reduced with respect to f_i.

Note that the $i < j$ condition is superfluous because automatically f_i is reduced with respect to f_j for $i < j$. Now, we try to make this more sense of this by introducing an ordering of polynomials:

Definition 7 (Ordering Polynomials) — We say that $f < g$ if either $\text{class}(f) < \text{class}(g)$ or $\text{class}(f) = \text{class}(g)$ and $\text{LD}(f) < \text{LD}(g)$. If neither $f < g$ and $g < f$ hold, then we write $f \sim g$.

We can show this is a well-ordering:

Proposition 3.1
$<$ is a well-ordering on $k[x_1, \ldots, x_n]$.

Proof. Consider some set S of polynomials. If $S \cap k$ is nonempty, that is a minimal element. Otherwise, consider $S' \subseteq S$ containing all polynomials of minimal class. Then, by the well-ordering of the integers, S' has an element of minimal degree, and by the definition of $<$, this is the minimal element of S as well. □

What about ascending chains? We can do the same thing with them:

Definition 8 (Ordering Chains) — Let $C_f = f_1, \ldots, f_r$ and $C_g = g_1, \ldots, g_s$ be two ascending chains. Then $C_f < C_g$ if there exists a t such that $f_i \sim g_i$ for $i < s$ and $f_s < g_s$, or $r > s$ and $f_i \sim g_i$ for all $1 \leq i \leq s$.

Not surprisingly, this is also well-ordered:

Proposition 3.2

$<$ is a well-ordering on the set of ascending chains.

Proof. Consider some subset Γ of ascending chains, and consider the subset Γ_1 be the subset of all ascending chains such that their first polynomial is minimal compared to any first polynomial. If there is a chain $p \in \Gamma_1$ with only one term, it is minimal; otherwise, we look at Γ_2 defined similarly. Repeating this process as many times as the length of the maximal chain, we get that for some s, there exists a chain C with s elements in Γ_s, and the first C to satisfy this property is the ascending chain. □

Now, we can use this definition very nicely:

Definition 9 (Characteristic Sets) — For any nonempty $S \subseteq k[x_1, \ldots, x_n]$, call the minimal ascending chain the characteristic set of S.

As we need to write algorithms, we can show that there exists a finite algorithm to find the characteristic set. However, we need the following propositions first:

Proposition 3.3

Consider the characteristic set $C = f_1, \ldots, f_r$ of S with $\text{class}(f_1) > 0$. If g is a nonzero polynomial reduced with respect to C, then the characteristic set C' of $S \cup \{g\}$ is less than C.

Proof. We shall find a smaller element. If $\text{class}(g) \leq \text{class}(f_1)$, as $g < f_1$, the set $\{g\}$ itself suffices, as that is less than C. Now, if $\text{class}(g) > \text{class}(f_1)$, and let j be the largest integer such that $\text{class}(g) > \text{class}(f_j)$. Then, note $C' = f_1, \ldots, f_j, g$ is an ascending chain as each is reduced with respect to all other polynomials. Note $C' < C$ as $f_k \sim f_k$ for all $1 \leq k \leq j$ and $g < f_{j+1}$. □

We can use this to fuel the following result:

Proposition 3.4 (Characterizing Characteristic Sets)

An ascending chain $C = f_1, \ldots, f_r$ with $\text{class}(f_1) > 0$ is a characteristic set of S if and only if there are no nonzero reduced polynomials in S (with respect to C).

Proof. If g was some nonzero reduced polynomial, apply the previous proposition. Now, suppose we had an ascending chain satisfying the constraints. Furthermore, suppose $C' = g_1, \ldots, g_s < C$ is a smaller ascending chain. If there exists some t such that $f_k \sim g_k$ for all $1 \leq k < t$ and $g_t < f_t$, but then g_t is reduced, a contradiction. Thus, we must have $r < m$, but then g_{r+1} is reduced, a contradiction. □

Now, we can prove the following theorem:

Theorem 3.5

Every nonempty S has a characteristic set, and finite S have a (finite) algorithm to find the characteristic set,

Proof. The first statement follows as the ascending chains are well-ordered. If $\text{class}(f) = 0$, then f_1 is the characteristic set. Consider $S_1 = \{g \in S \mid g \text{ is reduced with respect to } f_1\}$ (which is done by computing $\deg(g, \text{LV}(f_1))$ for all $g \in S$). Now, S_1 only consists of 0 polynomials, we are done. Otherwise, let f_2 be the minimal element of S_1 and let $S_2 = \{g \in S \mid g \text{ is reduced with respect to } f_2\}$. Continuing like this, we can get that if $\{f_1, \ldots, f_r\}$ is a characteristic set (we can not go on infinitely as S is finite). □

4. Ritt's Principle

What if we didn't consider the polynomial set, but the ideal generated by that set? That's the notion behind the following definition:

Definition 10 (Extended Characteristic Sets) — Consider some subset $S = \{h_1, \ldots, h_m\} \subseteq k[x_1, \ldots, x_n]$ and I be the ideal generated by them. The ascending characteristic set is an ascending chain $C = f_1, \ldots, f_r$ such that either $r = 1$ and $f_1 \in k \cap I$ or $f_i \in I$ and $\text{pr}(h_j, f_1, \ldots, f_r) = 0$.

This definition won't make sense until we show the following proposition:

Proposition 4.1

Consider some subset $S = \{h_1, \ldots, h_m\} \subseteq k[x_1, \ldots, x_n]$ and I be the ideal generated by them. If $C = f_1, \ldots, f_r$ is the extended characteristic set of S, then it is the characteristic set of I.

Proof. If $r = 1$, then I doesn't contain any nonzero reduced elements, so applying Characterizing Characteristic Sets finishes the job. Otherwise, we proceed by contradiction, assuming $g \in I$ is reduced. Now, as $\text{pr}(h_j, f_1, \ldots, f_r) = 0$, we can write

$$h_j \prod_{i=1}^{r} d_i^{s_i} = \sum_{i=1}^{r} Q_{i,j} f_i$$

Now, we can note that

$$g = \sum_{i=1}^{m} A_j h_j$$

for some polynomials A_j, so thus we can say that

$$g \prod_{i=1}^{r} d_i^{s_i} = \sum_{j=1}^{m} A_j \sum_{i=1}^{r} Q_{i,j} f_i = \sum_{i=1}^{r} f_i \left(\sum_{j=1}^{m} A_j Q(i,j) \right)$$

contradicting the fact that it was nonzero. □

Now, just like characteristic sets, we can also show that extended characteristic sets have an algorithm to find them:

Theorem 4.2 (Ritt's Principle)

Consider some finite subset $S = \{h_1, \ldots, h_m\} \subseteq k[x_1, \ldots, x_n]$ and I be the ideal generated by them. There exists a (finite) algorithm to find the extended characteristic set C of S.

Proof. Note that we can construct the characteristic C_1 of $S = S_1$. If this is only a constant, this is our desired characteristic set. Otherwise we can expand S_1 to S_2 such that for every $g \in S$, $\text{pr}(g, f_1, \ldots, f_r) \in S_2$ (where $C_1 = f_1, \ldots, f_r$). If $S_1 = S_2$, that means that $\text{pr}(h_j, f_1, \ldots, f_r) = 0$, so thus C_2 (the characteristic set of S_2) is the desired characteristic set.

Continuing to repeat this process, we can get an increasing chain

$$S_1 \subseteq S_2 \subseteq \cdots$$

and a decreasing chain

$$C_1 > C_2 > \cdots$$

Thus, either $S_j = S_{j-1}$ for some j or C_j is a constant (by the well-ordering of subsets and chains), which implies that we have found the characteristic as C_j. However, we still need to show $f_i \in I$, which will follow if we can induct to show $C_i \subset S_i \subset I$.

Base Case. $i = 1$. In this case, we need to show $C \subset S \subset I$, which follows by their definitions.

Induction Hypothesis. Suppose for $i \geq 2$, $C_{i-1} \subset S_{i-1} \subset I$.

Induction Step. Now, we note that S_i is just adding on the remainders upon pseudodivision, but as the remainder lies in I, we must have $S_i \subset I$. As $C_i \subset S_i$, we are done. □

A lot of what we did related to using the triangular form. However, when we could factor different things, we may return into degenerate conditions of the problem statement. Thus, we try to look at irreducibility in polynomial chains.

Definition 11 (Irreducible Ascending Chains) — Let $C = f_1, \dots, f_r$ be an ascending chain in $k[x_1, \dots, x_n]$. Renaming some variables so that we can write

$$f_j \in k[u_1, \dots, u_d, x_1, \dots, x_k]$$

we say C is irreducible if f_i is irreducible in $k(u_1, \dots, u_d)[x_1, \dots, x_i]/(f_1, \dots, f_i)$.

Here we will note that we assume that x_i is the leading variable of f_i in order to suppress hidden conditions. Now, here's a theorem that tries to use pseudodivision to help us relate with irreducible ascending chains:

Theorem 5.1

Let $C = f_1, \dots, f_r$ be an irreducible ascending chain and $g \in k[u_1, \dots, u_d, x_1, \dots, x_r]$ and $F_r = k(u_1, \dots, u_d)[x_1, \dots, x_d]/$
Then, the following four statements are equivalent:

(1) $\operatorname{pr}(g, f_1, \dots, f_r) = 0$
(2) Let E be any extension of the field k. If $\mu = (\tilde{u_1}, \dots, \tilde{u_d}, \tilde{x_1}, \dots, \tilde{x_r}) \in E^{d+r}$ is in $V(f_1, \dots, f_r)$ with $\tilde{u_1}, \dots, \tilde{u_d}$ transcendental in k, then $\mu \in V(g)$.
(3) g is 0 as an element of F_r.
(4) There exist finitely many nonzero polynomials $c_1, \dots, c_s \in k[u_1, \dots, u_d]$ with $c_1 \cdots c_s g \in (f_1, \dots, f_r)$.

However, to show that 1 implies 2, we will need the following lemma:

Lemma 5.2

Let p be a polynomial in x_m (for $1 \leq m \leq r$) with degree $s \geq 0$ and coefficients as polynomials in $k[u_1, \dots, u_d, x_1, \dots, x_{m-1}]$. Suppose p was reduced with respect to $f_1, \dots, f_r$. If μ is a zero of p, then p is equivalently zero,

Proof of Lemma. Note that if $s = 0$, given that a constant polynomial has a root, it must be the zero polynomial. We can assume $s \geq 1$ and shall induct on m:
Base Case. $m = 1$. Let $\tilde{q}$ be $q(\mu)$. Then, we get

$$\tilde{p} = \sum_{j=0}^{s} \tilde{a_j} \tilde{x_1}^j$$

As p is reduced with respect to f_1, we have $s < \deg(f_1, x_1)$. Now, suppose we evaluated $f(\tilde{u_1}, \dots, \tilde{u_d}, x_1)$. This is irreducible in $k(\tilde{u_1}, \dots, \tilde{u_d})[x_1]$, so thus $\tilde{p}$ is in the field extension

$$k(\tilde{u_1}, \dots, \tilde{u_d})(x_1) \cong k(\tilde{u_1}, \dots, \tilde{u_d})[x_1]/(f_1(\tilde{u_1}, \dots, \tilde{u_d}, x_1))$$

However, as in this extension $\tilde{p}$ is the zero polynomial, we get $\tilde{a_j} = 0$. However, as each of the $\tilde{u_j}$ were transcendental, $a_j = 0$, so thus $p = 0$.
Induction Hypothesis. Assume the result is true for some $m - 1 \geq 1$. We shall demonstrate the result for m.
Induction Step. We repeat the same process. As p is reduced with respect to f_m, we have $s < \deg(f_m, x_m)$. Now, suppose we evaluated $f(\tilde{u_1}, \dots, \tilde{u_d}, x_1)$. This is irreducible in $k(\tilde{u_1}, \dots, \tilde{u_d})[x_1, \dots, x_m]$, so thus $\tilde{p}$ is in the field extension

$$k(\tilde{u_1}, \dots, \tilde{u_d})(x_1, \dots, x_m) \cong k(\tilde{u_1}, \dots, \tilde{u_d})[x_1, \dots, x_m]/(f_1(\tilde{u_1}, \dots, \tilde{u_d}, x_1, ldots, x_m))$$

However, as in this extension $\tilde{p}$ is the zero polynomial, we get $\tilde{a_j} = 0$. Applying our induction hypothesis shows $a_j = 0$, which means $p = 0$, completing the proof. □

Now, we can move on to the proof of the theorem:

Proof of Theorem. First we show 2 implies 1. Suppose $g(\mu) = 0$ and let $R = \text{pr}(g, f_1, \ldots, f_r)$. We know we have

$$d_1^{s_1} \cdots d_r^{s_r} g = \sum_{j=1}^{r} A_j f_j + R$$

Note that $f_j(\mu) = 0$, so thus $R = 0$ by our lemma.

Now we show 1 implies 2. We note

$$d_1^{s_1} \cdots d_r^{s_r} g = \sum_{j=1}^{r} A_j f_j$$

As by the definition of an ascending chain, we know $\text{pr}(d_j, f_1, \ldots, f_r)$ is nonzero, which means that $d_k(\mu) \neq 0$. However, this directly implies that $g(\mu) = 0$.

Note that 3 and 2 directly imply each other by moving over all possible values of μ.

Now, we shall show 4 implies 2. As $\tilde{u_i}$ is transcendental, we get $c_i(\mu) \neq 0$. In particular, we get $\mu \in V(f_1, \ldots, f_r)$ and $g(\mu) = 0$, which is the statement in 2.

Now, we wrap up by showing 1 implies 4. Suppose we could write

$$d_1^{s_1} \cdots d_r^{s_r} g = \sum_{j=1}^{r} A_j f_j$$

First, we note that if $d_1^{s_1} \cdots d_r^{s_r}$ was zero, we could write it as

$$d_1^{s_1} \cdots d_r^{s_r} = \sum_{j=1}^{r} Q_j f_j$$

contradicting the assumption that d_j was reduced. Thus, $d_1^{s_1} \cdots d_r^{s_r}$ has an inverse in F_r, so we can write

$$qp - 1 = \sum_{j=1}^{r} Q_j f_j$$

or by clearing denominators we get

$$q_1 p - c = \sum_{j=1}^{r} \overline{Q_j} f_j$$

However, we also know

$$q_1 \left(\sum_{j=1}^{r} A_j f_j \right) = d_1^{s_1} \cdots d_r^{s_r} g q_1 = p g q_1 = g \left(\sum_{j=1}^{r} \overline{Q_j} f_j + c \right)$$

Thus, gc is in the ideal, completing the proof. □

References

[1] Joran Elias. https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1034&context=tme. Automated Geometric Theorem Proving: Wu's Method.

[2] Wu's method for automated geometry theorem proving and discovering. https://www.sciencedirect.com/science/article/pii/B9780127347608500065?via%3Dihub