A Refutational Approach to Geometry Theorem Proving*

Deepak Kapur**

General Electric Company, Corporate Research and Development, Schenectady, NY 12345, U.S.A.

ABSTRACT

A refutational method for proving universally quantified formulae in algebraic geometry is proposed. A geometry statement to be proved is usually stated as a finite set of hypotheses implying a conclusion. A hypothesis is either a polynomial equation expressing a geometric relation or a polynomial inequation (the negation of a polynomial equation) expressing a subsidiary condition that rules out degenerate cases and perhaps some general cases. A conclusion is a polynomial equation expressing a geometry relation to be derived. Instead of showing that the conclusion directly follows from the hypothesis equations and inequations, the proof-by-contradiction technique is employed. It is checked whether the negation of the conclusion is inconsistent with the hypotheses. This can be done by converting the hypotheses and the negation of the conclusion into a finite set of polynomial equations and checking that they do not have a common solution. There exist many methods for this check, thus giving a complete decision procedure for such geometry statements. This approach has been recently employed to automatically prove a number of interesting theorems in plane Euclidean geometry. A Gröbner basis algorithm is used to check whether a finite set of polynomial equations does not have a solution. Two other formulations of geometry problems are also discussed and complete methods for solving them using the Gröbner basis computations are given.

1. Introduction

In the early days of Artificial Intelligence research, there was considerable interest in developing computer programs for automatically proving geometry theorems. Gelehrter and his colleagues [17] investigated automating geometric reasoning using a synthetic approach to geometry. Mimicking techniques in

* A preliminary version of this paper appeared in Proceedings ACM Symbolic and Algebraic Computation (SYMSAC) Conference held in July 1986 at Waterloo, Canada. Some of the results reported in this paper appeared without proofs as an application letter in the Journal of Symbolic Computation.

** Current address: Department of Computer Science, State University of New York at Albany, Albany, NY 12222, U.S.A.
high school geometry books, Gelernter characterized various geometric concepts including parallel line, vertical angle, collinear, congruence, equilateral triangle, etc., by axioms and rules of inference. These axioms and rules of inference along with geometric construction techniques and examples were used to prove some simple plane geometry theorems; see also [18]. This work was subsequently extended by Nevins using a combination of forward and backward chaining methods. Coelho and Pereira [15] recently surveyed this approach toward geometry theorem proving and discussed its limitations.

Much before Gelernter's efforts, there was a major breakthrough in algebra, geometry and logic when Tarski [39] showed in the 1930s that the theory of elementary algebra and elementary geometry is decidable. Subsequently, Tarski's decision algorithm, which used an ingenious quantifier elimination technique, was improved by others—notably among them Seidenberg [37], Monk [34], and Collins [16], and recently by Ben-Or et al. [4]. In contrast to Gelernter's synthetic method, these methods are algebraic and are based on translating geometry statements into first-order formulae using the operations 0, 1, −1, +, *, ÷, = of an ordered field with variables ranging over the real numbers. Among these decision procedures, Collins' method based on the cylindrical algebraic decomposition technique is, to our knowledge, the only decision procedure implemented so far; see [2, 3] for details. Unfortunately, Collins' method has been found too general and slow to be useful in proving plane geometry theorems.

Recently, Wu Wen-tsun [42, 43] has revived interest in automated geometry theorem proving by showing how a subclass of geometry theorems (expressed using the concepts of parallelism, incidence and congruence) can be proved using a fairly simple and elegant algebraic method. A geometry statement of the form—a finite set of hypotheses implying a conclusion—is considered. Hypotheses are polynomial equations expressing geometric relations and the conclusion is also a polynomial equation stating a geometric relation to be derived; a subset of variables corresponding to coordinates of points that can be arbitrarily chosen in a geometry statement are viewed as parameters. Unlike the methods of Tarski, Seidenberg, Monk, Collins and Ben-Or et al., variables in Wu's method range over an algebraically closed field.

This method has been implemented by Wu in China and Chou at the University of Texas independently. Both Wu [42, 43] and Chou [10, 11] have extensively experimented with the method, and have been successful in proving many nontrivial theorems in geometry. Chou's program [10], in particular, has proved over 200 theorems in Euclidean and non-Euclidean geometries. Further, Wu's method has also been used for proving some interesting results about perspective projections in the application area of image understanding [25, 35, 38].

In this paper, an alternative approach for proving similar algebraic geometry theorems is proposed. Two types of geometry problems are considered and complete decision methods are discussed for these problems. The first type
GEOMETRY THEOREM PROVING

consists of deciding whether a given universally quantified formula is a theorem or not. The second type of geometry problems (suggested by Wu's approach) includes the first type and consists of three parts:

(i) decide whether a given universally quantified formula is a theorem or not,

(ii) if it is not a theorem, check whether there is an additional hypothesis possibly ruling out degenerate cases and perhaps some general cases (called nondegenerate conditions or subsidiary conditions) such that the original formula subject to this additional hypothesis becomes a theorem, and

(iii) in case such a hypothesis exists, find it.

The proposed approach is based on Hilbert's Nullstellensatz and is complete in Wu's geometry but is incomplete in Tarski's geometry, that is, geometry statements are not restricted to hold over the real numbers only. Both Wu's approach and the proposed approach are sound in the sense that a geometry statement declared a theorem by these approaches is indeed a theorem even when interpreted over the real numbers only; however, the converse is not true. These approaches are thus good heuristics for deciding geometry statements over the reals.

Unlike Wu's approach, the proposed approach does not use factorization of polynomials over successive extension fields of a base field, which is generally considered a difficult problem (cf. [10, pp. 278–279; 46, p. 3]). Further, degenerate cases, which should be an integral part of a complete geometry statement, are handled in a natural way in this approach. It was suggested in a preliminary version of this paper [28] that degenerate cases can be stated a priori in Wu's approach also; recently, Wu [45] has suggested a method which can be used for handling degenerate cases, and Ko [30] has investigated its use in geometry theorem proving over the reals.

In the proposed approach, a geometry statement is translated to checking whether a finite set of polynomials does not have a common zero in an algebraically closed field. Since it follows from Hilbert's Nullstellensatz that a set of polynomials has a common zero if and only if their ideal is not the unit ideal, to prove a geometry statement it is sufficient to check whether an equivalent finite set of polynomials generates the unit ideal. There exist many methods for checking whether a given set of polynomials generates the unit ideal [6, 8, 9, 33, 36, 40, 41, 45]. In this paper, the use of the Gröbner basis method [5–8, 22] is investigated for this check. Kutzler and Stiftler [31, 32] and Chou and Schelter [12] have also independently investigated the use of the Gröbner basis method for proving geometry theorems. Their approaches are somewhat different; their works are discussed later in the paper.

1.1. Organization of the paper

The paper is organized as follows. In the next section, we briefly discuss Wu's method; we introduce the language of geometry statements under considera-
tion and call it Wu's geometry. In Section 3, two types of geometry problems are precisely defined. The fourth section relates theorem proving in algebraic geometry to Hilbert's Nullstellensatz. It is shown that deciding universally quantified geometry statements is equivalent to checking whether a finite set of polynomials corresponding to the geometry statement does not have a common zero. From this observation, one obtains a complete method for deciding universally quantified geometry statements discussed in Section 2. The incompleteness of the method as well as of Wu's method in Tarski's geometry of reals are shown even for this restricted class of geometry statements. Section 5 lists methods which could be used for checking whether a finite set of polynomial equations has no solution.

Section 6 discusses the use of Gröbner bases for geometry theorem proving. This method is illustrated on simple examples. A table of geometry statements proved by this method and the time taken on a Symbolics 3640 LISP machine for proving these theorems is given. Related work on using the Gröbner basis computation for proving geometry theorems is also discussed. In Section 7, it is shown how the proposed approach can also be used for deriving subsidiary conditions if a given geometry statement is not a theorem. The method is shown to be complete for deriving such conditions when interpreted over an algebraically closed field. Conditions deduced by this method are usually simpler and weaker than those reported in [10, 11] as well as in [12, 32, 33]. Tables giving the time taken on a Symbolics 3640 LISP machine to deduce subsidiary conditions for a number of geometry statements are given. Section 8 outlines how additional information given about certain points in a geometry statement being general can be used in deciding geometry statements. Some empirical observations made while using the proposed approach in proving a large class of geometry theorems are reported in Section 9. Section 10 discusses some differences between the methods based on Gröbner basis computations and Wu's method based on Ritt's characteristic sets.

2. Wu's Geometry

As stated above, Wu's method as well as the proposed method are algebraic. A coordinate system is associated with a geometry problem with points in a geometry problem having coordinates taken from an algebraically closed field. Further, geometry problems under consideration can only be stated using the concepts of parallelism, incidence and congruence.

Consider a field $k$ of characteristic zero intrinsic to geometry under consideration, for example, the field of rationals; let $K$ be an algebraically closed field containing $k$ [41]. Let $+, *, 0, 1, -1$ be the operations of $K$. Assume a denumerable set of variables ranging over $K$. Consider the set of all quantifier-free formulae constructed using these operations, the equality predicate and Boolean connectives, in which the variables range over $K$. Let this language be
W, after Wu, and such a geometry Wu’s geometry. It is worthwhile noting again that unlike Tarski’s language of elementary geometry, (i) there is no order relation on polynomials in this language; thus, geometry problems expressed using the betweenness relation cannot be stated in this language, (ii) alternation of quantifiers is not allowed, and (iii) variables do not range over the real closed field, instead they range over the algebraically closed field. Such geometries are called metric geometries in [13]. Thus, the language W is a highly restricted subset of Tarski’s language of elementary geometry.

2.1. Wu’s method

A geometry statement considered by Wu [42, 44, 46] is of the following type:

Given a finite set of hypotheses, say \( \{h_1, \ldots, h_i\} \), expressed as polynomials over \( k \), involving variables \( x_1, \ldots, x_n \), decide whether a given conclusion polynomial \( c \) is generically true or not.

Variables above are coordinates of points in a geometry statement. Variables corresponding to coordinates of points that can be “arbitrarily chosen” in a geometry statement are called parameters (or independent variables) by Wu whereas the remaining indeterminates are called dependent variables. Polynomials \( h_1, \ldots, h_i \) specify geometry relations, such as points being collinear, lines being parallel or perpendicular, etc., in a geometry statement; polynomial \( c \) expresses a geometry relation being derived such as two line segments being of the same length.

Wu’s method for deciding the above class of geometry statements is based on (i) generating from the hypothesis polynomials, a set of polynomials in a triangular form with certain additional properties, called a characteristic set, a concept due to Ritt (also called a basic set by Wu), (ii) checking whether the conclusion polynomial \( c \) pseudo-divides to 0 using the polynomials in the characteristic set. Characteristic sets are computed using pseudodivision by introducing a total ordering on dependent variables. For any given finite set of polynomials, its characteristic set exists and can be effectively computed. Wu also gave an algorithm for computing a characteristic set from a given finite set of polynomials.

Let \( \text{Zeros}(\{f_1, \ldots, f_i\}) \) be the set of all common zeros in \( K \) of \( f_1, \ldots, f_i \); this set is also called the algebraic variety or algebraic manifold defined by \( \{f_1, \ldots, f_i\} \). Wu [44, Theorem (Ritt), p. 230] showed that

\[
\text{Zeros}(\text{CHS}/J) \subseteq \text{Zeros}(\text{Hyp}) \subseteq \text{Zeros}(\text{CHS}) ,
\]

\[
\text{Zeros}(\text{Hyp}) = \text{Zeros}(\text{CHS}/J) \cup \bigcup_j \text{Zeros}(\text{Hyp}_j) ,
\]

where \( \text{CHS} \) is a characteristic set of \( \text{Hyp} = \{h_1, \ldots, h_i\} \), \( J \) is the product of
initials\(^1\) of polynomials in \(CHS\), and \(CHS\) is the enlarged set of polynomials obtained by adjoining \(I_j\), the initial of the \(j\)th polynomial in \(CHS\), to \(Hyp\). The set \(Zeros(CHS/J)\) is the set of zeros of \(CHS\) that are not zeros of the polynomial \(J\), i.e.,

\[ Zeros(CHS/J) = Zeros(CHS) - Zero(J). \]

In case the conclusion \(c\) does not pseudodivide to 0 with respect to \(CHS\), the geometry statement can be said to be not generically true only if the characteristic set \(CHS\) is irreducible [44, Theorem 2, p. 242]. Otherwise, in case a characteristic set is reducible, then the characteristic set must be decomposed using factorization over extension fields by successively adjoining the factors of polynomials in the characteristic set. Wu defined that a geometry statement is generically true if and only if the conclusion \(c\) pseudodivides to 0 on each of the irreducible basic sets obtained after decomposition.

A by-product of Wu’s method is the generation of a set of polynomials, \(\{s_1, \ldots, s_j\}\) representing degenerate cases such that the conclusion polynomial \(c\) vanishes at the zeros in \(K\) of \(\{h_1, \ldots, h_i\}\) that are not zeros of \(\{s_1, \ldots, s_j\}\). These polynomials are the initials of polynomials from irreducible basic sets used to pseudodivide \(c\) to 0. Negations of these polynomial equations, i.e., \(s_1 \neq 0, \ldots, s_j \neq 0\), or equivalently \(s_1 \cdots s_j \neq 0\), are called nondegenerate or subsidiary conditions. This is an important advantage of Wu’s approach to geometry theorem proving. One does not need to worry about ruling out degenerate cases, such as points not being distinct, circles being of zero radius, etc., in geometry problems, such degenerate cases are often tedious to precisely enumerate. An interested reader may consult [10, 42, 44, 46] for further details.

Chou [10, 11] experimented with Wu’s method and, based on his extensive implementation experience with Wu’s method, Chou developed a way to interpret the subsidiary conditions geometrically. An interested reader may consult [10, 42, 44, 46].

Wu [44] showed that his method is complete for deciding geometry theorems in a generic sense over an algebraically closed field. However, the major stumbling block in his method is the need to factor polynomials over successive extension fields of a base field in obtaining irreducible basic set(s) of hypothesis polynomials. Chou developed an algorithm for factoring polynomials in which indeterminates have at most degree two [10, 11]; he has found his factoring algorithm to be quite adequate and efficient for proving plane geometry theorems. However, factoring arbitrary polynomials over successive extension fields is generally considered a difficult problem (cf. [10, pp. 278–279; 46, p.]

\(^1\)The initial of a polynomial is its leading coefficient when the polynomial is expressed in recursive form as a polynomial in its leading variable; see [44]. The leading variable of a polynomial is the highest dependent variable (with respect to a chosen ordering) appearing in the polynomial.
3]. As a result, theorems which need factorization of arbitrary polynomials
over successive extension fields in Wu’s method are likely to be not easily
handled.

3. Geometry Problems in Wu’s Geometry

Two different problems in Wu’s geometry are considered. The first problem is
the decision problem of universally quantified formulae, which is, given a
universally quantified formula in Wu’s geometry, decide whether the formula is
a theorem or not. Usually, one is interested in deciding universally quantified
formulae having the following structure:

$$\forall x_1, \ldots, x_n \in K,$$

$$[[h_1 = 0 \text{ AND } \ldots \text{ AND } h_i = 0 \text{ AND } s_1 \neq 0 \text{ AND } \ldots \text{ AND } s_j \neq 0] \Rightarrow [c = 0]].$$

Each of the $h, s$ and $c$ are in $k[x_1, \ldots, x_n]$. As in the case of Wu’s method,
polynomials $h_1, \ldots, h_i$ specify geometry relations, such as two lines being
parallel, a line intersecting a circle, etc.; $c$ expresses a geometry relation being
derived, such as two line segments being of equal length.

Polynomials $s_1, \ldots, s_j$ correspond to (i) degenerate cases such as points
being distinct, line segments being of nonzero length, etc., (ii) complex zeros
of the hypotheses, and (iii) general subcases of the hypotheses on which the
conclusion may not hold. Case (iii) often arises in the form of the reducibility
condition in Wu’s method in which, as stated above, the characteristic set
obtained from the hypotheses is not irreducible and, further, the conclusion
does not hold on all irreducible components obtained after factorization but
only on some components. Wu [46] and Chou and Yang [13] have recently
discussed such geometry problems. Wu’s method as discussed in [42, 44, 46]
did not allow polynomial inequations as hypotheses. However, the structure
theorem presented by Wu in [45] allowed polynomial inequations in the
hypotheses also.

It is easy to see that the above decision problem is the same as checking
whether the zeros in $K$ of $c$ include all the zeros of $\{h_1, \ldots, h_i\}$, on which
$s_1, \ldots, \text{ and } s_j$ do not vanish, i.e., whether

$$\text{Zeros}(\{h_1, \ldots, h_i\}/s_1, \ldots, s_j) \subseteq \text{Zeros}(\{c\}).$$

The second problem is more general than the first problem and is suggested by
Wu’s approach for handling geometry statements; this author is not aware of
any other literature in the field of theorem proving in which this problem has
appeared in such form. Its formulation is:

Given a universally quantified formula, decide whether it is a
theorem or not, and if it is not a theorem, change the formula
“slightly” so that the modified formula is a theorem.
This formulation can be made precise in case of Horn formulae:

Given a finite set of hypotheses and a conclusion, decide whether the conclusion follows from the hypotheses and if not, find an additional hypothesis, if any, that is consistent with the original set of hypotheses such that the conclusion follows from the original set of hypotheses and this additional hypothesis.

Thus, we have:

Given a set of hypotheses \( \{h_1 = 0, \ldots, h_i = 0\} \), and a conclusion \( c = 0 \), if the hypotheses are consistent and \( c \not\in k \), and \( m \leq n \), decide whether \( c = 0 \) holds and if not, find \( s_i \), if any, expressed in \( \{x_1, \ldots, x_m\} \) such that

(i) \( \forall x_1, \ldots, x_n \in K, \quad [[h_1 = 0 \ \text{AND} \ldots \ \text{AND} \ h_i = 0] \Rightarrow [s_1 = 0 \ \text{OR} \ldots \ \text{OR} \ s_j = 0]] \), is not a theorem, and

(ii) \( \forall x_1, \ldots, x_n \in K, \quad [[h_1 = 0 \ \text{AND} \ldots \ \text{AND} \ h_i = 0 \ \text{AND} \ s_1 \neq 0 \ \text{AND} \ldots \ \text{AND} \ s_j \neq 0] \Rightarrow [c = 0]] \), is a theorem.

The condition (i) ensures that the deduced hypothesis is consistent with the original set of hypotheses in a geometry statement. In the above, \( s_1, \ldots, s_j \) can be combined into a single polynomial \( s_1 \ast \cdots \ast s_j \) also. It is possible to consider a more flexible formulation in which the set of hypotheses could also include some subsidiary conditions expressed as polynomial inequalities.

It will of course be nice to put some additional restrictions on the \( s_i \) above so that they have a geometric meaning but it is not clear how that can be done. One possible way is to identify a subset of variables as parameters, as is done in Wu's method, and require that the \( s_i \) be polynomials in these parameters. Later in the paper, it is discussed how such information can be used for solving the above two geometry problems.

The first geometry problem is considered next in Sections 4, 5, and 6. In Section 7, the second geometry problem is considered. Complete decision methods are given for both problems.

4. Hilbert's Nullstellensatz and Algebraic Geometry

It is shown how a geometry statement involving Boolean connectives including negations can be reduced with the help of new variables to a conjunction of polynomial equations.

Definitions. A polynomial equation \( p = 0 \), where \( p \in k[x_1, \ldots, x_n] \), is satisfiable (or consistent, \( p = 0 \) has a solution, or equivalently, \( p \) has a zero in \( K \)) if and only if there exist \( v_1, \ldots, v_n \) in \( K \) such that \( p(x_1 \leftarrow v_1, \ldots, x_n \leftarrow v_n) \), the result of substituting \( v_1, \ldots, v_n \) for \( x_1, \ldots, x_n \), respectively, in \( p \), also written...
as \( p(v_1, \ldots, v_n) \), evaluates to 0. Then, \((v_1, \ldots, v_n)\) is a zero of \(p\) in \(K^n\). An equation \(p = 0\) is unsatisfiable (or inconsistent) otherwise. Similarly, a polynomial inequation \(p \neq 0\) is satisfiable if and only if there exist \(v_1, \ldots, v_n\) in \(K\) such that \(p(v_1, \ldots, v_n)\) does not evaluate to 0; \(p \neq 0\) is unsatisfiable otherwise. Given a set \(S\) of polynomial equations and inequations in \(k[x_1, \ldots, x_n]\), \(S\) is satisfiable (or consistent or equivalently, \(S\) has a common solution) if and only if there exist \(v_1, \ldots, v_n\) in \(K\) such that for every polynomial equation \(p = 0\) in \(S\), \(p(v_1, \ldots, v_n)\) evaluates to 0, and for every polynomial inequation \(p \neq 0\), \(p(v_1, \ldots, v_n)\) does not evaluate to 0. The set \(S\) is unsatisfiable (or inconsistent) otherwise.

For deciding geometry statements, the proof-by-contradiction technique is employed in this paper, much like the refutational approach used in resolution-based theorem proving and equational approaches to theorem proving in predicate calculus [20, 26]. It is checked whether the hypotheses including the subsidiary conditions, and the negation of the conclusion together are unsatisfiable. Note that unlike Wu's formulation, there is no need to classify variables appearing in a geometry statement into independent variables and dependent variables.

In general, the problem is to decide whether a universally quantified formula in language \(W\) is a theorem or equivalently, whether an existentially quantified formula in \(W\) is unsatisfiable. It is shown below how the satisfiability problem of any arbitrary quantifier-free formula in Wu's geometry is equivalent to the problem of finding whether an equivalent finite set of polynomial equations is satisfiable. This is done by showing how to simulate various Boolean connectives by introducing new variables [37].

**Lemma 4.1.** The satisfiability of \(p \neq 0\) is equivalent to the satisfiability of \(pz - 1 = 0\), where \(z\) is a new variable.

This construction is similar to the one used in the proof of Hilbert's Nullstellensatz [41].

**Proof.** (\(\Rightarrow\)) There is an assignment of variables in \(p\) at which \(p\) does not vanish; then that assignment extended with \(z = 1/v\), where \(v\) is the value of \(p\) for that assignment, is a zero of \(pz - 1\).

(\(\Leftarrow\)) There is an assignment of variables at which \(pz - 1\) vanishes. Then \(p\) cannot vanish at that assignment. \(\Box\)

Similarly, we have:

**Lemma 4.2.** The satisfiability of \((p_1 = 0\text{ OR } p_2 = 0)\) is equivalent to the satisfiability of \(p_1 p_2 = 0\), and the satisfiability of \((p_1 = 0\text{ AND } p_2 = 0)\) is equivalent to the satisfiability of \(\{ p_1 = 0, p_2 = 0 \} \).
If there is more than one conclusion in a geometry statement, they can also be handled easily using the above transformations. From these transformations, it is obvious that the satisfiability of any quantifier-free formula involving Boolean connectives is equivalent to the satisfiability of an equivalent finite set of polynomial equations as constructed above. (For instance, convert an arbitrary quantifier-free formula into a conjunctive normal form, and then use the above transformations to obtain a set of polynomial equations.)

Theorem 4.3. The satisfiability of any quantifier-free formula in the language $W$ is equivalent to the satisfiability of an equivalent finite set of polynomial equations obtained as discussed above.

4.1. A decision procedure for universally quantified formulae of $W$

Using the above theorem, the problem of deciding the satisfiability of a quantifier-free formula in $W$ is equivalent to finding whether there exists a solution in $K$ of a finite set of equations. This is a special case of the problem addressed by Hilbert’s Nullstellensatz over $K$ [41, Vol. II, p. 157].

Hilbert’s Nullstellensatz. If $f$ is a polynomial of $k[x_1, \ldots, x_n]$ which vanishes at all the common zeros of $f_1, \ldots, f_i$ in the $n$-dimensional affine space of $K$, then

$$f^q = h_1 f_1 + h_2 f_2 + \cdots + h_i f_i$$

for some natural number $q$ and polynomials $h_1, h_2, \ldots, h_i \in k[x_1, \ldots, x_n]$.

A special case of the above theorem is the following corollary:

Corollary 4.4. If $f_1, \ldots, f_i$ have no common zero in the $n$-dimensional affine space of $K$, then 1 is in the ideal generated by $f_1, \ldots, f_i$, written as $1 \in (f_1, \ldots, f_i)$ [41, Vol. II, p. 157].

Theorem 4.5. The validity of a universally quantified formula in $W$ is equivalent to (i) finding whether an equivalent set of polynomial equations has no solution, and (ii) whether the ideal generated by the equivalent set of polynomials is the unit ideal.

As a corollary of the above theorem, we have:

Corollary 4.6. The validity of a geometry statement of the form:

$$\forall x_1, \ldots, x_n \in K, \quad \text{[[} h_1 = 0 \text{ AND } \ldots \text{ AND } h_i = 0 \text{ AND } s_i \neq 0 \text{ AND } \ldots \text{ AND } s_j \neq 0 \text{]} \Rightarrow [c = 0]$$
is equivalent to deciding whether the ideal \((h_1, \ldots, h_i, s_1z_1 - 1, \ldots, s_jz_j - 1, czz - 1)\) is the unit ideal, where \(h_1, \ldots, h_i, s_1, \ldots, s_j, c \in k[x_1, \ldots, x_n]\), and \(z_1, \ldots, z_j, zz\) are distinct variables different from \(x_1, \ldots, x_n\).

The problem of checking whether an ideal generated by a finite set of polynomials is the unit ideal or equivalently whether a finite set of polynomials has a common zero in an algebraically closed field is known to be decidable. This gives us a complete decision procedure for universally quantified formulae in Wu's geometry. Section 5 addresses the problem of checking whether a given finite set of polynomials generates the unit ideal.

4.2. Incompleteness of Wu's method and proposed method in Tarski's geometry

As the following example provided by Singer\(^2\) illustrates, both Wu's method and the proposed approach based on Hilbert's Nullstellensatz are incomplete in Tarski's geometry even for the decision problem of universally quantified geometry statements. Both methods consider complex zeros in contrast to Tarski's method which considers real zeros (see also [29]).

\[(\forall x)(\forall y)x^2 + y^2 = 0 \Rightarrow [x = 0 \text{ AND } y = 0].\]

The above formula is a theorem if \(x\) and \(y\) are assumed to range over the real numbers. However, the above is not a theorem if \(x\) and \(y\) are assumed to range over \(K\). If Wu's method is used, the hypothesis polynomial is already in triangular form (assuming \(x > y\)), and the conclusion polynomials cannot be pseudodivided. The above formula is not a theorem by the approach proposed in this paper either, as the polynomial equations

\[x^2 + y^2 = 0,\]
\[(xz_1 - 1)(yz_2 - 1) = 0,\]

do have a complex solution; for example take \(y^2 = -c^2\) for any nonzero rational \(c\); then \(x = c\) and \(y = ci\), \(z_1 = 1/c\), \(z_2 = d\) is a zero. However, these polynomial equations do not have a real solution.

5. Checking the Existence of a Solution of a System of Equations

There are many methods for deciding whether a finite set of polynomials generates the unit ideal or equivalently whether a finite set of polynomial equations has a solution; this is called the triviality problem of polynomial

\(^2\text{Personal communication, 1984.}\)
ideals in [1]. In this paper, the use of the Gröbner basis method [5–8, 22] is investigated for this problem because of the author's familiarity with the method. Other methods worth investigating are

(i) Hermann's method [19],
(ii) a method based on combining elimination, resultants and S-polynomials of the Gröbner basis computation [36],
(iii) Wu's method based on triangulation [45], and
(iv) a recent method proposed by Chistov and Grigoryev [9] which extends Lazard's method [33] and has a complexity of subexponential time.

Below some observations regarding Wu's method for this problem are discussed. In the next section, the use of the Gröbner basis method is illustrated.

5.1. Factoring is essential in Wu's method

If Wu's method is to be used for deciding whether an ideal generated by a set of polynomials is the unit ideal, factorization of polynomials over successive extension fields of a base field is essential in the triangulation procedure. The need for factorization in Wu's method is illustrated by the following simple example due to Narendran.\(^3\)

\[ p_1 = x^2 + 2x + 1, \quad p_2 = (x + 1)y^2 + x. \]

It is easy to see that the ideal generated by \(p_1\) and \(p_2\) is the unit ideal; this can be tested also using the Gröbner basis method. However, the two polynomials are in triangular form under the ordering \(x < y\). This shows that it is not just sufficient to bring the polynomials in triangular form to check whether a set of polynomials generates the unit ideal. The polynomial \(p_1\) can be factored, i.e., \(p_1 = (x + 1)^2\). So, if the problem is decomposed considering each of the factors (in this case, the two factors are identical), one gets \(p'_1 = (x + 1)\). The polynomial \(p'_1\) simplifies \(p_2\), giving 1 as the triangular form.

6. A Gröbner Basis Method for Geometry Theorem Proving

The concept of a Gröbner basis was introduced by Buchberger [5, 6] for deciding the ideal membership problem for polynomial ideals over a field, and other related problems. Since then, this concept has been extensively studied and extended to other algebraic structures; see [8, 22] for details. In a Gröbner basis algorithm, a polynomial is viewed as a rule for simplifying other polynomials. Informally, a Gröbner basis of an ideal is defined as a special basis such that every polynomial in the ideal simplifies to 0 using this basis. In particular,

\(^3\)Personal communication, 1985.
Theorem (Buchberger). A Gröbner basis of the unit ideal must include 1.

So, the decision problem for universally quantified formulae in Wu’s geometry can be done using the Gröbner basis computation. We now give a method used for deciding geometry theorems of the following kind:

\[
\forall x_1, \ldots, x_n \in K, \\
[[h_1 = 0 \text{ AND } \ldots \text{ AND } h_i = 0 \text{ AND } s_1 \neq 0 \text{ AND } \ldots \text{ AND } s_j \neq 0] \\
\Rightarrow [c = 0]].
\]

Method I. Given hypotheses \( h_1, \ldots, h_i, \) degenerate cases \( s_1, \ldots, s_j, \) and a conclusion \( c, \) each a polynomial in \( Q[x_1, \ldots, x_n]: \)

\[
1 \in \text{Gröbner}(\{ h_1, \ldots, h_i, s_1z_1 - 1, \ldots, s_jz_j - 1, czz - 1 \}, \\
Q[x_1, \ldots, x_n, z_1, \ldots, z_j, zz])?
\]

yes: return THEOREM

no: return FALSIFIABLE

end.

The function Gröbner above takes a finite set of polynomials and a polynomial ring from which these polynomials are taken as arguments, and computes a Gröbner basis of the ideal specified by the input basis. The reader may wish to consult [7, 8, 22] for an algorithm for computing a Gröbner basis as well as for testing membership of a polynomial in an ideal using its Gröbner basis.

The above method is illustrated on three simple examples which can be done by hand. The first example is due to Mundy; the second example is taken from [10]; the third example is from [11].

Example 6.1. Given three lines such that \( AB \) is perpendicular to \( AC \) and \( CD \) is perpendicular to \( AC, \) prove that \( AB \) is parallel to \( CD; \) see Fig. 1. Set

\[
A = (0, 0), \quad B = (x_1, y_1), \quad C = (x_3, y_3), \quad D = (x_2, y_2).
\]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>D</td>
</tr>
</tbody>
</table>

Fig. 1. A simple example.
The hypothesis polynomial equations are:
\[ y_1 y_3 = -x_1 x_3 \quad ; AB \text{ is perpendicular to } AC , \]  
(1)
\[ y_3 (y_2 - y_3) = -(x_2 - x_3)x_3 \quad ; CD \text{ is perpendicular to } AC . \]  
(2)

The subsidiary condition is that points \( A \) and \( C \) be distinct, i.e.,
\[ x_3 \neq 0 \text{ OR } y_3 \neq 0 . \]

By introducing new variables \( z_1 \) and \( z_2 \), the above transforms to
\[ (x_3 z_1 - 1)(y_3 z_2 - 1) = 0 . \]  
(3)

Polynomial equation for the conclusion is
\[ y_1 (x_2 - x_3) = x_1 (y_2 - y_3) \quad ; AB \text{ and } CD \text{ are parallel} . \]  
(4)

And, its negation translates, after introducing a new variable \( z z_1 \), to
\[ (y_1 (x_2 - x_3) - x_1 (y_2 - y_3))z z_1 - 1 = 0 . \]  
(5)

The Gröbner basis of the set \( \{1, 2, 3, 5\} \) of polynomials includes 1, which implies that under the condition that \( A \) and \( C \) are distinct points, the theorem holds. Some of the steps in the derivation of the Gröbner basis are given below; for details, the reader can consult [8]. The ordering on variables is
\[ x_3 < y_3 < x_1 < y_1 < x_2 < y_2 < z_1 < z_2 < z z_1 ; \]

the head-monomial of a polynomial is picked using the degree ordering (comparing terms by degree first and terms of equal degree by lexicographic (dictionary) ordering). The rules corresponding to the above polynomials are:
\[ y_3 y_1 \rightarrow -x_3 x_1 , \]  
(1)
\[ y_3 y_2 \rightarrow -x_3 x_2 + x_3 x_3 + y_3 y_3 , \]  
(2)
\[ x_3 y_3 z z_1 \rightarrow y_3 z_2 + x_3 z_1 - 1 , \]  
(3)
\[ x_1 y_2 z z_1 \rightarrow y_1 x_2 z z_1 - x_3 y_1 z z_1 + y_3 x_1 z z_1 - 1 . \]  
(4)

The overlapping of rules (2) and (4) gives a superposition \( x_1 y_2 y_3 z z_1 \) obtained by the least common multiple of the headterms on which both the rules can be applied as shown in Fig. 2; this gives a critical pair possibly resulting in a new rule.
GEOMETRY THEOREM PROVING

\[ x_1, y_2, y_3, zz_1 \]

\[ 2 \quad 4 \]

\[ -x_3 x_1, x_2, zz_1 + x_3 x_2, x_1, zz_1 + y_3 y_1, x_1, zz_1 \]

\[ y_3 y_1, x_2, zz_1 - y_3 x_3, y_1, zz_1 + y_3 y_3, x_1, zz_1 - y_3 \]

Fig. 2. Superposition and critical pairs.

After these polynomials are further simplified and their difference is taken, the polynomial \( y_3 \) is generated. The rule corresponding to this polynomial is:

\[ y_3 \rightarrow 0, \quad (5) \]

which simplifies all the above rules to:

\[ x_3 x_1 \rightarrow 0, \quad (1') \]

\[ x_3 x_2 \rightarrow x_3 x_3, \quad (2') \]

\[ x_3 z_1 \rightarrow 1, \quad (3') \]

\[ x_1 y_2 zz_1 \rightarrow y_1 y_2 zz_1 - x_3 y_1 zz_1 - 1. \quad (4') \]

Polynomial (1') now corresponds to \( x_3 \) being 0 or \( x_1 \) being 0. Since polynomial (3') corresponds to \( x_3 \) being not 0, rules (1') and (3') superpose to generate a new rule, \( x_1 \rightarrow 0 \). This new rule simplifies rule (4') to give

\[ y_1 y_2 zz_1 \rightarrow x_3 y_1 zz_1 + 1. \quad (4'') \]

This rule with rule (2') has a superposition \( x_3 y_1 x_2 zz_1 \), on which (2') and (4'') are applied and the resulting polynomials are further simplified. The difference of the resulting polynomials gives a new rule, \( x_3 \rightarrow 0 \), which with the rule (3') gives a rule \( 1 \rightarrow 0 \), a contradiction. Hence the proof.

**Example 6.2.** Let \( ABCD \) be a parallelogram as shown in Fig. 3; prove that \( |AO| = |OC| \).

![Fig. 3. Parallelogram example.](image)
The coordinate system chosen is
\[ A = (0, 0), \quad B = (u_1, 0), \]
\[ C = (u_2, u_3), \quad D = (x_2, x_1), \quad O = (x_3, x_4). \]

Polynomial equations of the hypotheses are:
\[ u_3 - x_1 = 0 \quad ; AB \text{ is parallel to } CD, \quad (6) \]
\[ x_1(u_2 - u_1) - x_2u_3 = 0 \quad ; BC \text{ is parallel to } AD, \quad (7) \]
\[ x_4(x_2 - u_1) - x_1(x_3 - u_1) = 0 \quad ; B, O, \text{ and } D \text{ are collinear}, \quad (8) \]
\[ x_3u_3 - x_4u_2 = 0 \quad ; A, O, \text{ and } C \text{ are collinear}. \quad (9) \]

The subsidiary conditions are: \( u_1 \neq 0 \text{ AND } u_3 \neq 0 \), which means that \( A, B, \) and \( C \) are distinct points which are not collinear. This condition translates to:
\[ u_1u_3x_1 - 1 = 0. \quad (10) \]

The polynomial equation for the conclusion is:
\[ (x_3^2 + x_4^2) = (u_2 - x_3)^2 + (u_3 - x_4)^2. \]

The polynomial equation corresponding to its negation is:
\[ zz_1((x_3^2 + x_4^2) - ((u_2 - x_3)^2 + (u_3 - x_4)^2)) - 1 = 0. \quad (11) \]

The Gröbner basis of the above polynomials is 1, implying that the above statement is a theorem.

**Example 6.3** (Theorem of orthocenter taken from [11]). Let \( ABC \) be a triangle, \( H \) be the point of intersection of altitudes \( BE \) and \( CF \). Show that \( AH \) is perpendicular to \( BC \). Let
\[ A = (0, 0), \quad B = (u_1, 0), \]
\[ C = (x_1, x_2), \quad E = (u_2, u_3), \quad H = (x_3, x_4). \]

Polynomial equations for the hypotheses are:
\[ x_2u_3 + (u_2 - u_1)x_1 = 0 \quad ; BE \text{ is perpendicular to } AC, \quad (12) \]
\[ -u_2x_2 + u_3x_1 = 0 \quad ; A, E, \text{ and } C \text{ are collinear}. \quad (13) \]
GEOMETRY THEOREM PROVING

Fig. 4. Theorem of orthocenter.

\[(u_1 - u_2)x_3 + u_3x_1 - u_1u_3 = 0 \; ; B, E, \text{ and } H \text{ are collinear}. \quad (14)\]

The subsidiary condition is \((x_2 \neq 0 \text{ AND } u_1 \neq 0)\) OR \(u_3 \neq 0\), which translates to

\[(z_1x_2u_1 - 1)(z_2u_3 - 1) = 0. \quad (15)\]

Polynomial equation for the conclusion is

\[x_2x_3 = -x_1(x_1 - u_1),\]

stating that \(AH\) is perpendicular to \(BC\). The negation of the conclusion is:

\[zz_1(x_2x_3 + x_1(x_1 - u_1)) - 1 = 0. \quad (16)\]

The Gröbner basis of these polynomials is 1 thus proving the original theorem.

In the above examples, subsidiary conditions are in the form of points being distinct or points being not collinear. Subsidiary conditions are also used to rule out certain general subcases of the hypotheses for which the conclusion does not hold, such as in the case of the secants theorem discussed in [32, 46] and [13, Example 3.3.4] discussed by Chou and Yang.

Method I above was used to prove a number of geometry theorems including Simson's theorem, Pascal's theorem, Pappus' theorem, Desargues' theorem, the nine-point circle theorem, the butterfly theorem, and Gauss' theorem. An implementation of the Gröbner basis algorithm for the polynomial rings over the integers and rationals developed by Rick Harris at General Electric Corporate Research and Development, was used for this purpose. This implementation incorporates optimizations for not considering certain critical pairs proposed by Buchberger [7] and Kapur et al. [24], and it supports lexicographic, degree, as well as mixed (see definition later) orderings on terms. The implementation runs on a Symbolics 3640 LISP machine. Table 1 gives computation times on some representative geometry theorems taken from [10–12, 31, 32]. The Gröbner basis computation for these theorems was performed using the degree ordering as that was found to be generally much
Table 1
Time needed to prove theorems

<table>
<thead>
<tr>
<th>Theorem</th>
<th>Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid</td>
<td>0.7</td>
</tr>
<tr>
<td>Ceva</td>
<td>20.9</td>
</tr>
<tr>
<td>Secants</td>
<td>9.6</td>
</tr>
<tr>
<td>Equidistant</td>
<td>0.2</td>
</tr>
<tr>
<td>Simson</td>
<td>13.7</td>
</tr>
<tr>
<td>Pappus</td>
<td>3.0</td>
</tr>
<tr>
<td>Square</td>
<td>0.3</td>
</tr>
<tr>
<td>Tangent circle</td>
<td>1.5</td>
</tr>
<tr>
<td>Peripheral angle</td>
<td>6.5</td>
</tr>
<tr>
<td>Altitudes</td>
<td>0.4</td>
</tr>
<tr>
<td>Desargues</td>
<td>1.2</td>
</tr>
<tr>
<td>Nine-point circle</td>
<td>12</td>
</tr>
<tr>
<td>Isosceles midpoint</td>
<td>6.4</td>
</tr>
<tr>
<td>Pentagon</td>
<td>0.05</td>
</tr>
<tr>
<td>Tetrahedron</td>
<td>5.2</td>
</tr>
<tr>
<td>Pappus’ dual</td>
<td>6.9</td>
</tr>
<tr>
<td>Quadrangle in $\mathbb{R}^2$</td>
<td>0.45</td>
</tr>
<tr>
<td>Quadrangle in $\mathbb{R}^3$</td>
<td>0.7</td>
</tr>
<tr>
<td>Gauss</td>
<td>0.2</td>
</tr>
<tr>
<td>Pascal</td>
<td>640</td>
</tr>
<tr>
<td>Brahmagupta</td>
<td>1.1</td>
</tr>
<tr>
<td>Wang</td>
<td>19.6</td>
</tr>
<tr>
<td>Butterfly</td>
<td>25.1</td>
</tr>
</tbody>
</table>

faster than the pure lexicographic ordering; see [8] for similar observations about the performance of the Gröbner basis algorithm using degree and lexicographic orderings on terms. As one can see from the table, most of the theorems can be automatically proved in less than a minute of computer time. These include fairly nontrivial theorems such as the butterfly theorem, the nine-point circle theorem, Simson’s theorem, Ceva’s theorem and Wang’s theorem, some of which humans find difficult to prove. To our knowledge, such theorems have not been proven using an automatic or semi-automatic theorem prover based on the synthetic approach initiated by Gelernter [15]. This shows the power of the algebraic methods. One serious drawback of algebraic methods though is that the proofs obtained cannot be translated back into a form that uses high-level geometry concepts so that they are understandable by humans.

6.1. Related work on the use of the Gröbner bases for geometry theorem proving

Kutzler and Stifter [31, 32] have independently investigated the use of the Gröbner basis method for geometry theorem proving. Their method assumes
that in a geometry statement, a subset of variables is identified to be parameters. This method is based on checking whether the conclusion polynomial is in the ideal generated by the hypothesis polynomials; they compute the Gröbner basis of polynomials over the field \( \mathbb{Q}(u_1, \ldots, u_d) \), where \( u_1, \ldots, u_d \) are parameters in Wu's sense. Polynomials in \( u_1, \ldots, u_d \), that are needed to be nonzero in checking whether the conclusion polynomial is in the ideal of hypothesis polynomials are produced as the subsidiary conditions by their method.

Chou and Schelter [12] have also been using a Gröbner basis of polynomial ideals over \( \mathbb{Q}(u_1, \ldots, u_d) \) for checking whether the conclusion polynomial is in the ideal of hypothesis polynomials under certain conditions. Further, instead of first generating a Gröbner basis and then checking whether the conclusion polynomial belongs to the ideal generated by the hypothesis polynomials, they make the check every time new polynomials are added to the basis. Their method also produces subsidiary conditions which are polynomials required to be nonzero in generating a Gröbner basis and simplifying the conclusion polynomial. They compared the Gröbner basis approach with Wu's method on 80 examples and found Wu's method to be faster than the Gröbner basis approach.

The approaches pursued by Kutzler and Stifter as well as Chou and Schelter are based on checking whether a conclusion polynomial is in the ideal generated by the hypothesis polynomials under subsidiary conditions. This is a first approximation to checking whether the conclusion polynomial is in the radical of the ideal generated by the hypothesis polynomials; however, as shown by their results, even this first approximation works on many examples. These approaches are incomplete. For completeness, if the conclusion polynomial is not in the ideal generated by the hypothesis polynomials, one needs to successively examine powers of the conclusion polynomial for membership. Chou and Schelter have also reported results based on an approach very similar to ours for geometry statements in which a subset of variables are identified as parameters, but they have not investigated it in detail.

Since the check for triviality of an ideal is a special case of the check for the ideal membership problem, the theoretical complexity of the ideal membership problem is obviously at least as much as the theoretical complexity of the triviality problem. In fact, we conjecture that the membership problem for the radical of a polynomial ideal is easier than the polynomial ideal membership problem. Further, if a polynomial \( p \) simplifies to 0 using a basis, then using the same basis and using at most one more reduction, the polynomial \( pz - 1 \), where \( z \) does not appear in \( p \), reduces to \( -1 \). Thus, a suitable implementation of the triviality check should be at least as fast as the membership check for ideals. This along with the fact that the method based on Hilbert's Nullstellensatz is complete for Wu's geometry are our reasons for investigating the proposed approach.
7. Deducing Subsidiary Conditions

In this section, the second geometry problem stated in Section 3 is discussed. Recall that the second geometry problem is:

Given a Horn formula, i.e., a finite set of hypotheses implying a conclusion, decide whether it is a theorem or not. If it is not a theorem, then find an additional hypothesis, if any, that is consistent with the original set of hypotheses such that with the additional hypothesis, the Horn formula is a theorem.

This problem is more general than the geometry problem considered by Wu and Chou as hypotheses could include polynomial inequations expressing subsidiary conditions to rule out some degenerate cases or general subcases. We show below how this problem can be solved using the Gröbner basis method. Although the proposed approach is inefficient as compared to Wu's approach, subsidiary conditions found using it are often simpler and weaker than the ones reported using Wu's method [10] or those reported in [12] and [31, 32] using the Gröbner basis approach. Before discussing the technical details, the method is illustrated using examples.

Example 6.1 Revisited. For instance, for Example 6.1 (Fig. 1), if a Gröbner basis of the hypothesis polynomials (1) and (2), and the polynomial (5), corresponding to the negation of the conclusion is computed without the subsidiary condition that points $A$ and $C$ be distinct, the result does not include 1. (In fact, the statement of Example 6.1 translates to polynomial equations (1), (2), and (5), as it does not mention that points $A$ and $C$ be distinct.) Instead, there appear $x_3$, $y_3$, and other polynomials in the Gröbner basis. This implies that the hypothesis polynomials indeed have common zeros which are not zeros of the conclusion polynomial; these zeros are $x_3 = 0$, $y_3 = 0$, and suitably chosen values of the remaining variables. Further, these zeros are real zeros, suggesting that the statement is a theorem only when such zeros are ruled out, i.e., when $x_3 \neq 0$ or $y_3 \neq 0$, which is the condition that $A$ and $C$ are distinct points thus making $AC$ indeed determine a line.

In contrast, if Wu's method is used for this example, we obtain stronger conditions. If we partition the variables such that $y_1$ and $y_2$ are dependent variables, then the hypothesis polynomials are already in triangular form. The conclusion polynomial can be pseudodivided to 0 under the condition that $y_3 \neq 0$. This condition is sufficient as imposing this condition would make points $A$ and $C$ distinct. However it is too strong. If instead, we choose $x_1$ and $x_2$ as dependent variables, then again, the hypothesis polynomials are still in triangular form with respect to this choice of variables. The conclusion polynomial can be pseudodivided to 0 under the condition that $x_3 \neq 0$, which is still too strong.
If any of the approaches in [12, 31, 32] based on computing a Gröbner basis of polynomial ideals over \( Q(u_1, \ldots, u_d) \) is used, the conditions obtained are too strong then also. For instance, the two hypothesis polynomials of Example 6.1 above constitute a Gröbner basis over \( Q(x_3, y_3, x_1, x_2) \). The conclusion polynomial (4) simplifies to 0 under the condition that \( y_3 \neq 0 \), which is stronger than the condition that points \( A \) and \( C \) be distinct. Similarly, the two hypothesis polynomials of Example 6.1 constitute a Gröbner basis over \( Q(x_3, y_3, y_1, y_2) \); the conclusion (4) simplifies to 0 with the condition that \( x_3 \neq 0 \).

**Example 6.2 Revisited.** Similarly, for Example 6.2 (Fig. 3), if a Gröbner basis is computed of the hypothesis polynomials (6), (7), (8) and (9), along with the polynomial (11) corresponding to the negation of the conclusion, and without the subsidiary condition that \( A, B, \) and \( C \) are distinct points which are not collinear, the result again does not include 1. Instead, the result has among other polynomials, \( u_1u_3 \). From this, we get the subsidiary condition that \( u_1u_3 \neq 0 \), which implies that \( A \) and \( B \) are distinct points and \( C \) is not collinear with \( AB \). However, if the hypotheses also include the subsidiary condition that \( A, B, \) and \( C \) are distinct points, computing the Gröbner basis still does not result in 1 in the basis; instead, the basis includes \( u_3 \), which implies that the theorem still does not hold in case \( u_3 = 0 \), i.e., when \( A, B, \) and \( C \) are distinct but collinear points.

**Example 6.3 Revisited.** For Example 6.3 (Fig. 4) also, the Gröbner basis of the hypothesis polynomials (12), (13) and (14), along with the polynomial (16) corresponding to the negation of the conclusion does not include 1; instead it has among other polynomials, \( x_2u_1 \) and \( u_3 \) in the basis. This gives us the subsidiary condition that \( (x_3 \neq 0 \text{ AND } u_1 \neq 0) \text{ OR } u_3 \neq 0 \), which states that \( A, B, \) and \( C \) are distinct and are not collinear.

For Pappus' theorem [10, Example 3], when a Gröbner basis is computed of the hypothesis polynomials along with the polynomial corresponding to the negation of the conclusion without the subsidiary conditions, the result includes

\[
x_1x_4, \ x_2x_5, \ x_3x_6, \ x_1x_2x_3, \ x_4x_5x_6
\]

as polynomials, among others. The subsidiary condition under which the theorem holds is a disjunction of the following conditions: (i) \( x_1x_4 \neq 0 \), i.e., points \( A_1 \) and \( B_1 \) are distinct, (ii) \( x_2x_5 \neq 0 \), i.e., \( A_2 \) and \( B_2 \) are distinct, (iii) \( x_3x_6 \neq 0 \), i.e., \( A_3 \) and \( B_3 \) are distinct, (iv) \( x_1x_2x_3 \neq 0 \), i.e., neither of \( A_1, A_2, \) and \( A_3 \) is the point of intersection of lines \( A_1A_2A_3 \) and \( B_1B_2B_3 \), (v) \( x_4x_5x_6 \neq 0 \), i.e., neither of \( B_1, B_2, \) and \( B_3 \) is the point of intersection of lines \( A_1A_2A_3 \) and \( B_1B_2B_3 \). For the square example, [10, Example 4], the condition found by our
method is simply that the square have a nonzero side. For the triangle altitudes theorem [10, Example 5], the subsidiary condition found using the Gröbner basis method is \( u_3 \neq 0 \).

Let us now precisely state the second problem as discussed in Section 3:

Given a consistent set of hypotheses \( \{ h_1 = 0, \ldots, h_i = 0 \} \), a conclusion \( c = 0 \) such that \( c \not\in k \), and \( m \leq n \), find \( s_i \), if any, expressed in \( \{ x_1, \ldots, x_m \} \) such that
(i) \( \forall x_1, \ldots, x_n \in K, [[ h_1 = 0 \ AND \ \ldots \ AND \ h_i = 0 ] \Rightarrow [ s_1 = 0 \ OR \ \ldots \ OR \ s_j = 0 ] ] \) is not a theorem, and
(ii) \( \forall x_1, \ldots, x_n \in K, [[ h_1 = 0 \ AND \ \ldots \ AND \ h_i = 0 \ AND \ s_1 \neq 0 \ AND \ \ldots \ AND \ s_j \neq 0 ] \Rightarrow [ c = 0 ] ] \) is a theorem.

It should be noted that there is no requirement above that the \( s_i \) should have a geometric meaning. It is not clear how such a requirement can be algebraically stated. Chou and Yang [13] have argued that if independent variables among the variables of a geometry problem are identified as a part of the problem statement and geometry statements are proved in a generic sense, then it is possible to impose a geometric requirement on the \( s_i \). Independent variables can be viewed as degrees of freedom in a geometry problem. In our opinion, classification of variables into independent and dependent variables is also purely syntactic and may not necessarily have any geometric meaning. Further, there is no guarantee that the \( s_i \) expressed as polynomials in independent variables will necessarily have any geometric meaning. At best, one can hope for an algebraic meaning. The above formulation indeed has a precise algebraic meaning.

If for a given geometry problem, a Gröbner basis of the ideal \( (h_1, \ldots, h_i, c x z z - 1) \) does not include 1, which implies that \( c \not\in \text{Radical}(h_1, \ldots, h_i) \), the radical ideal of \( (h_1, \ldots, h_i) \), then any polynomial \( p \in k[x_1, \ldots, x_m] \) from the Gröbner basis such that \( p \) is not a consequence of the hypotheses (i.e., \( p \not\in \text{Radical}(h_1, \ldots, h_i) \)), is a candidate to be used for stating an additional hypothesis. One can examine the Gröbner basis for this purpose and pick a desired "simplest" such candidate (in fact, a set of polynomials).

The following theorem serves as the basis of this approach (Method II below):

**Theorem 7.1.** Let \( \{ h_1 = 0, \ldots, h_i = 0 \} \) be a consistent set of hypotheses, and \( c = 0 \), where \( c \not\in k \), be a conclusion, such that there is a polynomial \( p \in k[x_1, \ldots, x_m] \), \( m \leq n \), and \( p \not\in \text{Radical}(h_1, \ldots, h_i) \) but \( p c \in \text{Radical}(h_1, \ldots, h_i) \). Let \( GB \) be a Gröbner basis of \( (h_1, \ldots, h_i, c x z z - 1) \) under a lexicographic ordering on terms induced by the ordering \( x_1 < \cdots < \).
$x_n < zz$, where $zz$ is an indeterminate different from $x_1, \ldots, x_n$. Then there exists a polynomial $q \in GB$ such that

(i) $q \in k[x_1, \ldots, x_n]$ and $q \not\in \text{Radical}(h_1, \ldots, h_i)$,

(ii) $s_1, \ldots, s_j$ are the irreducible factors of $q$, and

(iii) $qc \in \text{Radical}(h_1, \ldots, h_i)$, thus implying

$$\forall x_1, \ldots, x_n \in K, \quad [\text{[} h_1 = 0 \text{ AND } \ldots \text{ AND } h_i = 0 \text{ AND } s_1 \neq 0 \text{ AND } \ldots \text{ AND } s_j \neq 0 \text{]} \Rightarrow [c = 0]].$$

**Proof.** Let $J = (h_1, \ldots, h_i, c zz - 1)$ and $J' = J \cap k[x_1, \ldots, x_n]$. Since there is a polynomial $p \in k[x_1, \ldots, x_n]$ such that $pc \in \text{Radical}(h_1, \ldots, h_i)$, for some $m$, there is an $r$ such that $(pc)^r \in (h_1, \ldots, h_i)$. It is easy to see that $p^r \in J'$ since

$$p^r = (pc zz)^r - p^r((c zz)^r - 1).$$

So, $p^r$ reduces to 0 using polynomials in $x_1, \ldots, x_m$ from $GB$. There is at least one polynomial among them which is not in $\text{Radical}(h_1, \ldots, h_i)$, as otherwise $p \in \text{Radical}(h_1, \ldots, h_i)$ leading to a contradiction. Call that $q$. The rest of the proof follows. □

The following method is a complete decision method for deriving subsidiary conditions. It is often better to identify independent variables and pick polynomials from $GB$ expressed in independent variables only to state subsidiary conditions. Examples reported in Tables 2 and 3 were computed in this manner.

**Method II.** Given hypotheses $h_1, \ldots, h_i$, and a conclusion $c$, each a polynomial in $Q[x_1, \ldots, x_n]$, and $m \leq n$:

\[
\{g_1, \ldots, g_r\} := \text{Gröbner}(\{h_1, \ldots, h_i, c zz - 1\}, Q[x_1, \ldots, x_n, zz]);
\]

if $g_1 = 1$
then return THEOREM: NO CONDITION NEEDED
else repeat from $v = 1$ to $r$
if $g_v \in Q[x_1, \ldots, x_m]$ and $g_v \not\in \{h_1, \ldots, h_i\}$
then if $1 \not\in \text{Gröbner}(\{h_1, \ldots, h_i, g_v zz - 1\}, Q[x_1, \ldots, x_n, zz])$
then return THEOREM UNDER CONDITION $g_v \neq 0$;
end repeat;
return THEOREM NOT CONFIRMED
end;
The function Gröbner above is assumed to return polynomials in a Gröbner basis in ascending order using the lexicographic ordering on terms. Notice that in the above method, the function Gröbner may be called several times. For geometry statements in which additional subsidiary conditions are required, two calls to the function Gröbner are often sufficient—the first for computing the Gröbner basis of \((h_1, \ldots, h_i, c \; zz - 1)\) and the second to ensure that the subsidiary condition deduced is consistent with the hypotheses. Picking an appropriate polynomial expressing the subsidiary conditions from the Gröbner basis of \((h_1, \ldots, h_i, c \; zz - 1)\) such that it has a geometric meaning is found to be quite a challenge. Method II returns THEOREM NOT CONFIRMED only in case the hypotheses and conclusion do not satisfy the conditions of Theorem 7.1. Because of a lack of any geometric justification of this method, a potential disadvantage is that this method may produce an additional hypothesis even if the conclusion \(c\) cannot be derived from the original set of hypotheses in the geometric sense.

Table 2 gives timings for deducing subsidiary conditions for some representa-

<table>
<thead>
<tr>
<th>Problem</th>
<th>First call</th>
<th>Second call</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid</td>
<td>1.2</td>
<td>0.4</td>
<td>Three corners of the triangle are not collinear</td>
</tr>
<tr>
<td>Ceva</td>
<td>312</td>
<td>5.3</td>
<td>Three corners of the triangle are not collinear</td>
</tr>
<tr>
<td>Secants</td>
<td>~3200</td>
<td>24.6</td>
<td>Well-defined secants</td>
</tr>
<tr>
<td>Equidistant secants</td>
<td>0.25</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Simson</td>
<td>~13100</td>
<td>1.2</td>
<td>Well-defined triangle, stated as sides being of nonzero length</td>
</tr>
<tr>
<td>Pappus</td>
<td>138</td>
<td>6.4</td>
<td>Two lines do not intersect at any of the given points</td>
</tr>
<tr>
<td>Square</td>
<td>0.9</td>
<td>0.2</td>
<td>Square is of nonzero size</td>
</tr>
<tr>
<td>Tangent circle</td>
<td>7.1</td>
<td>10.3</td>
<td>Secant is not a tangent</td>
</tr>
<tr>
<td>Peripheral angle</td>
<td>6.6</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Altitudes</td>
<td>8.2</td>
<td>4</td>
<td>Two corners of the triangle are distinct</td>
</tr>
<tr>
<td>Desargues</td>
<td>13.4</td>
<td>3.4</td>
<td>Two lines do not coincide</td>
</tr>
<tr>
<td>Nine-point circle</td>
<td>309</td>
<td>7.1</td>
<td>Triangle is not right-angled and has a side of nonzero length</td>
</tr>
<tr>
<td>Isosceles midpoint</td>
<td>738</td>
<td>7.0</td>
<td>Well-defined triangle with (D) distinct from (A) (cf. [12, Example 8])</td>
</tr>
<tr>
<td>Pentagon</td>
<td>0.08</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Tetrahedron</td>
<td>6.8</td>
<td>2.2</td>
<td>Tetrahedron is well-defined</td>
</tr>
<tr>
<td>Dual of Pappus</td>
<td>~21100</td>
<td>1750</td>
<td>Three points are distinct</td>
</tr>
<tr>
<td>Quadrangle in (R^2)</td>
<td>0.35</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Quadrangle in (R^3)</td>
<td>0.15</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Gauss</td>
<td>0.17</td>
<td></td>
<td>No condition needed</td>
</tr>
<tr>
<td>Pascal</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brahmagupta</td>
<td>446</td>
<td>41.3</td>
<td>Circle of nonzero radius and 4 points are distinct ((u_1 - u_i)(u_1(u_2 - u_i) - u_i(u_2 - u_i)) \neq 0) (cf. [12, Example 10])</td>
</tr>
<tr>
<td>Wang</td>
<td>2456</td>
<td>990</td>
<td></td>
</tr>
<tr>
<td>Butterfly</td>
<td>?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
tive geometry problems along with an English description of the subsidiary condition. The first column gives timings for the first call to the function Gröbner; the second column gives timings for the second call to Gröbner for checking that the deduced subsidiary condition is indeed consistent with the hypotheses.

From empirical observations, it is known that computing a Gröbner basis is often much more time consuming using a lexicographic ordering on terms than the degree ordering. The following refinement of Theorem 7.1 above is perhaps the best we can hope for in which degree and lexicographic orderings are combined; we will call such an ordering on terms a mixed ordering.

Given a total ordering on variables, say $x_1 < \cdots < x_n$, another total term ordering mixing the degree ordering and lexicographic ordering with $x_i$ as the separating variable is defined as:

$$t_1 = x_1^{d_1} \cdots x_n^{d_n} \leq_m t_2 = x_1^{g_1} \cdots x_n^{g_n}$$

if and only if

$$x_1^{d_1} \cdots x_n^{d_n} <_d x_1^{g_1} \cdots x_n^{g_n}$$

OR $(x_1^{d_1} \cdots x_n^{d_n} = x_1^{g_1} \cdots x_n^{g_n}$ AND $x_1^{d_1} \cdots x_{i-1}^{d_{i-1}} <_d x_1^{g_1} \cdots x_{i-1}^{g_{i-1}})$,

where $<_d$ is the degree ordering on terms.

Note that if the separating variable is chosen to be $x_1$, then the mixed ordering is the same as the degree ordering. The above total ordering on terms induces a total ordering on polynomials in a natural way as is done in [8]. Trinks [40] considered such a class of orderings also. The following theorem is a generalization of Theorem 7.1.

**Theorem 7.2.** Let $\{h_1 = 0, \ldots, h_i = 0\}$ be a consistent set of hypotheses, and $c = 0$, where $c \not\in k$, be a conclusion, such that there is a polynomial $p \in k[x_1, \ldots, x_m]$, $m \leq n$, and $p \not\in \text{Radical}(h_1, \ldots, h_i)$ but $pc \in \text{Radical}(h_1, \ldots, h_i)$. Let $GB$ be a Gröbner basis of $(h_1, \ldots, h_i, c \bar{z}z - 1)$ under a mixed ordering on terms induced by the ordering $x_1 < \cdots < x_{n+1} = \bar{z}z$, where $x_{m+1}$ is the separating variable and $\bar{z}z$ is an indeterminate different from $x_1, \ldots, x_n$. Then there exists a polynomial $q \in GB$ such that

(i) $q \in k[x_1, \ldots, x_m]$ and $q \not\in \text{Radical}(h_1, \ldots, h_i)$,

(ii) $s_1, \ldots, s_j$ are the irreducible factors of $q$, and

(iii) $qc \in \text{Radical}(h_1, \ldots, h_i)$, thus implying

$$\forall x_1, \ldots, x_n \in K$$

$$[[h_1 = 0 \text{ AND } \ldots \text{ AND } h_i = 0 \text{ AND } s_1 \neq 0 \text{ AND } \ldots \text{ AND } s_j \neq 0]$$

$$\Rightarrow [c = 0]].$$
Table 3
Statistics for deducing subsidiary conditions

<table>
<thead>
<tr>
<th>Problem</th>
<th>Mixed ordering</th>
<th>Degree ordering</th>
<th>Checking consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid</td>
<td>1.5</td>
<td>1.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Ceva</td>
<td>172</td>
<td>150</td>
<td>10.3</td>
</tr>
<tr>
<td>Secants</td>
<td>289</td>
<td>215</td>
<td>96</td>
</tr>
<tr>
<td>Equidistant secants</td>
<td>0.3</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Simson</td>
<td>723</td>
<td>492</td>
<td>336</td>
</tr>
<tr>
<td>Pappus</td>
<td>28.5</td>
<td>11.3</td>
<td>21</td>
</tr>
<tr>
<td>Square</td>
<td>1.2</td>
<td>0.75</td>
<td>0.2</td>
</tr>
<tr>
<td>Tangent circle</td>
<td>6.6</td>
<td>3.4</td>
<td>5.1</td>
</tr>
<tr>
<td>Peripheral angle</td>
<td>9.3</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>Altitudes</td>
<td>2.6</td>
<td>2.0</td>
<td>2.7</td>
</tr>
<tr>
<td>Desargues</td>
<td>6.6</td>
<td>4.7</td>
<td>3.2</td>
</tr>
<tr>
<td>None-point circle</td>
<td>225</td>
<td>164</td>
<td>48.6</td>
</tr>
<tr>
<td>Isosceles midpoint</td>
<td>192</td>
<td>124</td>
<td>107.6</td>
</tr>
<tr>
<td>Pentagon</td>
<td>0.08</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Tetrahedron</td>
<td>9.3</td>
<td>5.6</td>
<td>2.2</td>
</tr>
<tr>
<td>Dual of Pappus</td>
<td>516</td>
<td>106</td>
<td>1169</td>
</tr>
<tr>
<td>Quadrangle in $\mathbb{R}^2$</td>
<td>0.15</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Quadrangle in $\mathbb{R}^3$</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>Gauss</td>
<td>0.2</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Pascal</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Brahmagupta</td>
<td>68</td>
<td>55</td>
<td>29.7</td>
</tr>
<tr>
<td>Wang</td>
<td>868</td>
<td>628</td>
<td>320</td>
</tr>
<tr>
<td>Butterfly</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

The proof of Theorem 7.2 is the same as that of Theorem 7.1. Table 3 gives the timings using the mixed ordering for computing Gröbner bases with $zz$, the variable introduced to obtain a polynomial equivalent to the negation of the conclusion, as the separating variable. For the second call computation to check whether the deduced subsidiary condition is indeed consistent with the hypotheses, timings in Table 3 are given when the computation was done using the degree ordering. For each example, the condition deduced in this case is the same as the one reported in Table 2; so they are omitted from this table. Timings using the degree ordering for computing the Gröbner bases are also included in Table 3 for comparison. In the case of the degree ordering, we have not been able to prove a result analogous to Theorems 7.1 and 7.2; we conjecture that in this case, the result does not hold. However, for all examples discussed above, a suitable polynomial in $GB$ computed using the degree ordering to state subsidiary conditions was found.

7.1. Picking subsidiary conditions

If a single polynomial is selected from a Gröbner basis for stating a subsidiary condition, then a subsidiary condition is a conjunction of polynomial inequa-
tions, each corresponding to an irreducible factor of the polynomial $g_\nu$ above. Method II can be easily modified to select in general a set of polynomials from a Gröbner basis instead of a single polynomial. In that case, a subsidiary condition is a disjunction of conjunctions of polynomial inequations. A subsidiary condition, in general, rules out three types of cases: (i) geometrically degenerate cases, (ii) common complex zeros of the hypotheses which are not zeros of the conclusion, and (iii) general cases for which the conclusion does not follow from the hypotheses. Developing a good method to distinguish between these three types is an interesting open research problem.

It was observed in practice that computing a Gröbner basis without a subsidiary condition can be quite time consuming on big examples. For all examples, deducing subsidiary conditions took more time than proving a geometry theorem when subsidiary conditions were stated as part of the input. For examples such as the butterfly theorem and Pascal's theorem, their Gröbner bases could not be computed without subsidiary conditions in a reasonable amount of time. In many cases, it is possible to derive subsidiary conditions without having to fully compute a Gröbner basis. It is possible to preempt the computation after it has "sufficiently progressed" (meaning that all the hypothesis polynomials have interacted with the negation of the conclusion), and examine the basis to pick polynomials for deriving subsidiary conditions.

8. Using Information That Some Variables Are Parameters

Wu and Chou have argued that an integral part of a geometry statement is the specification of general points in a geometry statement. As Chou has shown, this information can be used to identify parameters (independent variables) among the variables appearing in a geometry statement. As discussed in Section 6.1, following this approach, Chou and Schelter as well as Kutzler and Stifter [31, 32] have investigated the use of Gröbner bases for geometry theorem proving by computing Gröbner bases of polynomial ideals over rational functions expressed in terms of these parameters. However, their approaches crucially depend upon precisely identifying these parameters among variables. Kutzler and Stifter have given a way of testing whether a subset of variables is correctly identified as parameters in a geometry statement using the concept of independence of variables.

Below, we outline how our approach can be used to exploit such information about a subset of variables to decide the two geometry problems discussed in Sections 6 and 7. It is not necessary to assume that all the parameters are identified in a geometry statement.

Let $u_1, \ldots, u_d$ be a subset of parameters among the variables $x_1, \ldots, x_n$ appearing in a geometry statement; let $\{y_1, \ldots, y_{n-d}\}$ be the remaining set of variables. Unlike in Section 3, where each variable ranges over $K$, in this case, every variable in $\{y_1, \ldots, y_{n-d}\}$ ranges over $K(u_1, \ldots, u_d)$. 
For the decision problem of universally quantified formulae, compute a Gröbner basis of the hypotheses polynomials with the negation of the conclusion polynomial and see whether 1 is generated during the computation. The Gröbner basis algorithm now computes over \( k(u_1, \ldots, u_d)[y_1, \ldots, y_{n-d}] \). If yes, then the conclusion follows from the hypotheses under the condition obtained by multiplying out the coefficients of all the polynomials that were used to reduce the polynomial corresponding to the negation of a conclusion. If 1 is not generated during the computation, then the conclusion does not follow from the hypotheses under the conditions which can be expressed as negations of polynomials in \( k[u_1, \ldots, u_d] \). Either the choice of parameters is incomplete or wrong, or the conclusion does not follow from the hypotheses.

For the second geometry problem, the method is the same as discussed in Section 7 except that the Gröbner basis algorithm now computes over \( k(u_1, \ldots, u_d)[y_1, \ldots, y_{n-d}] \). If 1 is not in a Gröbner basis, pick a polynomial(s) in \( u_1, \ldots, u_d, y_1, \ldots, y_i \) from the Gröbner basis that can be used to express an additional hypothesis. Another alternative is to modify the set of parameters.

9. Empirical Observations

We list below some observations that we made while trying to prove geometry theorems using the methods discussed in this paper. Some of these observations are similar to the ones made for generating canonical systems for equational theories using the rewriting approach as well as for computing Gröbner bases of polynomial ideals. However, some are quite specific to the use of the Gröbner basis approach for geometry theorem proving. Some of these observations are also similar to the ones made by Chou about Wu's method [10].

(1) The choice of an ordering on variables significantly affects the performance of this method. The dependence of the method on this choice is, however, not as sensitive as in Wu's method where one ordering may lead to an irreducible characteristic set, while another may lead to a reducible characteristic set thus requiring decomposition of the problem into many subproblems. (Chou claimed to have solved this problem of choosing an appropriate ordering for Wu's method.) Further, in contrast to Wu's method and Chou's method as well as the approaches discussed in [12, 31, 32], there is no need to classify variables into independent variables and dependent variables. It is not always clear how to make this selection. In cases where such a choice can be made (even partially), it may be computationally better to view a geometry problem as involving polynomials whose coefficients are rational functions over independent variables as discussed in Section 8. It may be easier to derive subsidiary conditions and missing hypotheses using this mixed approach.

(2) The order in which critical pairs are generated in the Gröbner basis
computation also considerably affects the performance. The incorporation of
the optimizations suggested in [7, 24], which significantly reduces the number
of critical pairs that must be considered, made significant improvements in the
performance of the Gröbner basis algorithm.

(3) The time taken to decide a geometry statement using this method
depends on the kind of subsidiary conditions in the statement. Including
additional subsidiary conditions in a statement does not necessarily reduce the
time taken to decide the statement.

(4) Different equivalent formulations of the same geometry problem (based
on choosing the origin and the axes, etc.) also significantly affect the timings.
Examples are the butterfly theorem and Simson's theorem.

(5) In checking whether a Gröbner basis of an ideal includes 1 (i.e., deciding
whether a geometry statement is a theorem), keeping the right-hand sides of
rules in bases in normal form does not affect the performance much; for some
examples it helps, while for other examples it slows the computation a little.
However, for deducing subsidiary conditions for geometry statements for
solving the second geometry problem, keeping the right-hand sides of these
rules in normal form considerably helps the performance while computing
Gröbner bases.

10. Comparison with Wu's Method

Chou and Schelter [12] have experimentally compared Wu's method with their
approach based on the Gröbner basis computation, and their results indicate
that Wu's method takes less time than the approaches based on the Gröbner
basis computation. Comparison of timings given in Table 1 with the timings for
Wu's method available from Chou for these examples also suggests that Wu's
method is faster than the proposed approach. Below, we give plausible reasons
why Wu's method takes less computational time than the approaches based on
the Gröbner basis computation; see also [14].

We feel that the pseudodivision and triangulation are strongly related to the
critical pair computation among polynomials which is central to the Gröbner
basis method. For instance, consider the four hypothesis polynomials in
Example 6.2. They are not in triangular form with respect to the ordering
$x_4 > x_3 > x_2 > x_1 > u_3 > u_2 > u_1$. In particular, polynomial (8) has to be
pseudodivided using polynomial (9) to get a triangular form. The result is:

$$(u_2x_2 - u_1u_2)x_3 + u_4u_2x_1 = 0.$$ 

(8')

If a critical pair is computed between polynomials (8) and (9), then one
obtains polynomial (8') above.

Critical pair computation is, however, significantly reduced in obtaining a
triangular form as compared to computing a Gröbner basis. In triangulation,
critical pairs are computed only among certain subset of polynomials; one of the polynomials in a critical pair computation is the smallest polynomial among all polynomials with the same highest variable. There is however a price to be paid, which is the need to ensure that every polynomial \( p \) introducing a new variable is irreducible over the extension field defined by the polynomials lower in the ordering than \( p \). As discussed in Section 5.1, this requirement is essential and a triangular form not satisfying this condition cannot be used to check for ideal membership.

Kandri-Rody, in his dissertation [21] (see also [23]), gave an algorithm for computing a different triangular form for a basis of an ideal which was used to develop a primality test of an ideal as well as to compute the dimension of an ideal. His triangular form, called an *extracted characteristic set* following Ritt, is obtained from a Gröbner basis using pure lexicographic ordering and thus does not have the limitation of triangular forms used by Wu and Chou. Kandri-Rody proved that every polynomial in an ideal pseudodivides to 0 using this triangular form ([21, Theorem 4.2.1, p. 68]).

Since for geometry theorem proving as well as for theorem proving in predicate calculus, we are interested only in the triviality problem, namely whether an ideal generated by a set of polynomials is the unit ideal, it will be worthwhile exploring a basis similar to a Gröbner basis, such that for the unit ideal, the basis includes 1 but for ideals that are not unit, it does not necessarily have the properties of a Gröbner basis. This is also related to the problem of generating term rewriting systems which compute canonical forms for certain equivalence classes of interest, and not necessarily canonical forms for all equivalence classes induced by a congruence relation.

11. Conclusion

We have discussed two complete decision methods for a subclass of geometry problems using a Gröbner basis approach. As shown above, these methods as well as Wu's method, which are algebraic, have been successful in proving many nontrivial plane geometry theorems. As far as we can tell, most of these theorems have not been proven using geometry theorem provers based on the synthetic approach [15]. Some of these theorems have been found extremely difficult to prove by humans including computer scientists. This seems to suggest that these algebraic methods have a lot of promise. Wu's method has also been used to prove nontrivial properties of perspective viewing in image understanding applications [25, 35, 38]; the reader may also see a paper by Kapur and Mundy in this volume.

One of the serious drawbacks of the algebraic methods is that the reasoning employed in these methods is significantly different from the reasoning used by humans while proving such theorems. Human reasoning in developing proofs of geometry problems is more along using synthetic properties and some
trigonometric properties. An interesting area of research is to develop methods and heuristics which translate algebraic manipulations used in Gröbner basis approaches as well as Wu's approach in geometry theorem proving to synthetic concepts so that an understandable proof can also be generated from these algebraic manipulations. This research will also be helpful in integrating algebraic methods with synthetics methods. Such integration of the two approaches will especially be useful in proving difficult theorems that may need human guidance. This is also likely to be helpful in the application of the algebraic methods in practical applications such as image understanding and solid modeling.

Note Added in Proof

The complexity of checking whether a finite set of polynomial equations has a solution has been shown to have an upper bound of single exponential time by Brownawell [47]. Further, it has been recently announced that such a bound can be achieved using a Gröbner basis algorithm (quoted in [48]). The worst-case complexity for computing a Gröbner basis is, however, double exponential time.

ACKNOWLEDGMENT

Rick Harris implemented the Gröbner basis algorithm on Symbolics LISP machine. David Cyrluk provided his lively company during my experiments and subsequently implemented a more tolerable interface to the Gröbner basis algorithm for geometry theorem proving. Shang-Ching Chou persuaded me during his visit to GE R&D Center in 1985 to experiment once again with the Gröbner basis approach. Hai-Ping Ko provided many geometry statements in polynomial forms so that I could quickly test them using my approach. Paliath Narendran provided a lot of ideas for the improvement of the method; many ideas in this paper originated during discussions with him. David Musser and Joe Mundy provided encouragement and support; thanks are especially to Joe who posed a challenge in the early summer of 1984 whether the Gröbner basis method and rewrite rules could be used for geometry theorem proving.

REFERENCES

5. Buchberger, B., An algorithm for finding a basis for the residue class ring of a zero-
dimensional polynomial ideal (in German), Ph.D. Thesis, University of Innsbruck, Austria (1965).
11. Chou, S.-C., Proving and discovering theorems in elementary geometry using Wu’s method, Ph.D. Thesis, Department of Mathematics, University of Texas, Austin, TX (1985).
GEOMETRY THEOREM PROVING


41. van der Waerden, B.L., Modern Algebra I, II (Fredrick Ungar, New York, 1966).


Received September 1986; revised version received January 1988