Journal of Automated Reasoning **00:** 1–20. 1999.

Automatic Discovery of Theorems in Elementary Geometry

T. RECIO
Dep. de Matemáticas, Univ. de Cantabria. 39071 Santander, Spain.
e-mail: recio@matesco.unican.es

M. P. VÉLEZ*
Dep. de Ingeniería Informática. Univ. Antonio de Nebrija. 28040 Madrid. Spain.
e-mail: pvelez@dii.unnet.es

(Received: 30 April 1996)

Abstract. We present here a further development of the well-known approach to automatic theorem proving in elementary geometry via algorithmic commutative algebra and algebraic geometry. Rather than confirming/refuting geometric statements (automatic proving) or finding geometric formulae holding among prescribed geometric magnitudes (automatic derivation), in this paper we consider (following Kapur and Mundy) the problem of dealing automatically with arbitrary geometric statements (i.e., theses that do not follow, in general, from the given hypotheses) aiming to find complementary hypotheses for the statements to become true. First we introduce some standard algebraic geometry notions in automatic proving, both for self-containment and in order to focus our own contribution. Then we present a rather successful but noncomplete method for automatic discovery that, roughly, proceeds adding the given conjectural thesis to the collection of hypotheses and then derives some special consequences from this new set of conditions. Several examples are discussed in detail.

Key words: automatic theorem proving, elementary geometry, Gröbner basis.

Introduction

We present here a further development of the automatic theorem proving method proposed by Kapur (see [6, 10]) after the dissemination of some foundational work by Wu [16]. Such an approach follows a rather straightforward algebraic geometry argument to refute elementary geometry theorems, and its algorithmic core is founded on Gröbner basis computations. Drawbacks and limitations of this method have been described in [1] and are taken into consideration here. The book [17] contains a detailed account of the whole subject by one of its founders, plus many references.

Roughly, automatic proving deals with deciding the correctness of statements of the kind $H \Rightarrow T$, where H, T are, respectively, some given sets of hypotheses

* Partially supported by DGICYT-PB 95-0354.

and theses. The expected output is, essentially, a yes/no answer, although it can happen that the method describes also some minor modifications of the hypotheses for the theses to hold (such as ruling out degenerate instances of the hypotheses). On the other hand, different methods for automatic proving have been adapted to the automatic derivation of geometric statements, where the goal is to find thesis involving prescribed geometric magnitudes that hold for a set of given hypotheses (such as deriving the expression of the area of a triangle in terms of the length of the three sides; see [2] and the recent work of [15] or [9]). We will consider in this paper a third issue, that of *automatic discovery*, that is, the problem of dealing automatically with arbitrary statements (i.e., statements that could be, in general, false) aiming to find complementary hypotheses for the statements to become true, for instance, stating that Brahmagupta's (Ptolomeus's) formula holds over arbitrary ordered quadrilaterals and getting as output that the vertices should be co-cyclic, an example taken from [15]. As has been already remarked, one application of this approach could be in image understanding, for example, to deduce conditions under which parallel lines on a scene remain parallel in its image (see [8]).

Automatic discovery is different from automatic proving in that the latter provides useful information only about the general (roughly speaking, see the discussion in Subsections 1.3 and 2.1) truth of a given statement. On the other hand automatic derivation is different from automatic discovery in that the former does not include a priori a specific thesis and does not pretend to modify the given hypotheses. Nevertheless, from the methodological point of view, automatic derivation could be understood as a particular case of automatic discovery of statements without a thesis at all (as the method presented in 2.1 could be easily adapted to derive geometric formulae), but we will not focus on this analogy here.

Although [8] already formulates explicitly the goals of automatic discovery* and proceeds with rather similar ideas (in the context of Wu's automatic proving techniques) to those presented here using Gröbner basis tools, we feel that most previous work just includes some enlightening examples and does not systematically address this important issue.

We have collected in Section 1 some needed notation and results from automatic proving, and we introduce one example that will be considered from different points of view along the paper. The reader is referred to book [1] or [4] for further details. Section 1 is written for the nonspecialist and has a didactic goal (the specialist should excuse our winding approximation to the key concept of geometrically independent variables, general components, and so on). Section 2 presents our automatic discovery method and discusses two illustrative examples. Some basic ideas have been already sketched in [14]. The book [13] (in Spanish) includes a relatively large collection of theorems (re)discovered under this approach. Finally, some conclusions are outlined, and a didactic application is briefly presented.

* ...the objective here is to find the missing hypotheses so that a given conclusion follows from a given incomplete set of hypotheses....

1. Automatic Proving

1.1. GEOMETRICALLY TRUE THEOREMS

Let K be a field of characteristic 0, for instance, the field of rational numbers $\mathbb{Q}$, and L an algebraically closed field containing K, for instance, the field of complex numbers $\mathbb{C}$. Given a geometric theorem (or, simply, a geometric construction), we begin our procedure by translating the geometric *hypotheses* and *theses* (or the construction steps) into algebraic expressions, after adopting a suitable coordinate system. Roughly speaking, the collection of hypotheses of a theorem is expressed as a set of polynomial equations*,

$$h_1(x_1, \ldots, x_n) = 0, \ldots, h_p(x_1, \ldots, x_n) = 0,$$

and the thesis** is also rewritten as a polynomial equation,

$$k(x_1, \ldots, x_n) = 0,$$

where $h_1, \ldots, h_p, k \in K[x_1, \ldots, x_n]$. Thus, the geometric statement is translated into

$$\forall (x_1, \ldots, x_n) \in L^n, h_1(x_1, \ldots, x_n) = 0, \ldots, h_p(x_1, \ldots, x_n) = 0$$
$$\Rightarrow k(x_1, \ldots, x_n) = 0.$$

Therefore, within this formulation, the geometric statement is said to be *geometrically true* iff the algebraic variety defined by $\{h_1 = 0, \ldots, h_p = 0\} \subset L^n$ (the hypothesis variety H) is contained in the variety $\{k = 0\} \subset L^n$ (the thesis variety T).

Since, by using the well-known trick of Rabinowistch, it happens that the ideal

$$\mathrm{J} = ((h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t]) \cap K[x_1, \ldots, x_n]$$

is also the set of elements $g \in K[x_1, \ldots, x_n]$ such that, for some power m, $gk^m \in (h_1, \ldots, h_p)$, it follows from Hilbert Nullstellensatz ‡ that a way for testing $H \subset T$ is to verify whether

$$1 \in \mathrm{I} = (h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t],$$

* It is quite easy to deal as well with hypotheses of the kind $h \neq 0$, by adding some extra variables z so that $h \neq 0$ is converted into $hz - 1 = 0$.

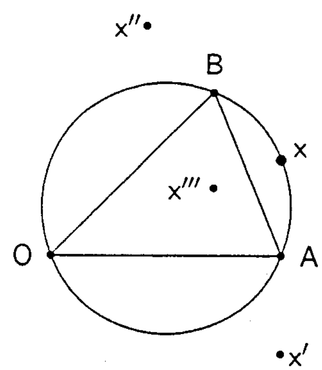
** If there is more than one, we can consider the conjunction of several theorems, one for each thesis.

‡ This formulation implies that we are proving/refuting statements taking into consideration all complex values of the involved variables. Although elementary Euclidean geometry, in general, deals with figures in the real plane or space, it happens that many "classical" theorems hold also when considering complex coordinates for the involved geometric constructions. The approach to automatic theorem proving or discovery of thesis holding only for real values of the coordinates is, computationally speaking, more difficult: see [3, 5, 12].

where t is some slack variable; and this can be automatically accomplished by computing any Gröbner basis of this ideal and by reporting if 1 is an element of such basis. Of course, in order to avoid proving trivially true results, it should be previously checked that H is nonempty (again, using the Nullstellensatz), because *any* thesis can be formally derived from a contradictory (empty) set of hypotheses.

There are many examples that show how the outlined procedure works (see the impressive collection given in [1]), but let us choose instead an example that does not work as intuitively expected.

EXAMPLE 1. *Let us consider a triangle of vertices* O(0, 0), A(l, 0), B(a, b) *and a point* x(x, y) *in the circle defined by the three given vertices (the outer circle or circumcircle of the triangle). We construct (see figure below) the symmetrical images of such point with respect to the three sides of the triangle, and we call them, respectively,* x′(x, $-y$), x″(X, Y), x‴(Z, W). *Now we claim that these three points are* always *aligned, that is,* $XW - Zy + xY + yX - YZ - xW = 0$.



To prove this theorem automatically we start generating the algebraic translation for the given data. Here we do it very naively – as it could be done by a machine with a minimum of information about the terms involved in our construction. For instance, the symmetric point x″ of x with respect to side $\overline{\text{OB}}$, is obtained step by step constructing, first, a perpendicular line to such side through x, namely: $a(L - x) + b(M - y) = 0$. Then, let (L, M) be the intersection of this line with the one described by the given side, so that also $-bL + aM = 0$. Finally, this symmetric point(X, Y) is defined through the vector equation $(x, y) + 2((L, M) - (x, y)) = (X, Y)$. For side $\overline{\text{AB}}$ we will define analogously an intermediate point (R, S) in the construction; and for the third side, the symmetric point can be directly seen that has coordinates $(x, -y)$. Next we consider the equation $x^2b - xbl + y^2b - ya^2 + yal - yb^2 = 0$, describing the circle that passes through O(0, 0), A(l, 0), B(a, b). In total, this yields the following collection of construction hypotheses:

$$\text{Ideal}\ (a(L - x) + b(M - y), -bL + aM, X - x - 2(L - x),\\ Y - y - 2(M - y), (a - l)(R - x) + b(S - y),\\ -b(R - l) + (a - l)S, Z - x - 2(R - x), W - y - 2(S - y),\\ x^2b - xbl + y^2b - ya^2 + yal - yb^2).$$

Obviously, this set of conditions is nonempty, for we are able to construct particular triangles verifying all of them. Next we test if the thesis follows from these hypotheses by checking (with the program **CoCoA*** all through this paper) the NormalForm** of 1 in I, where

$$\begin{aligned} \mathsf{I} = \ \mathsf{Ideal}(&La + Mb - xa - yb, -Lb + Ma, -2L + X + x, \\ &-2M + Y + y, Ra - Rl + Sb - xa + xl - yb, -Rb + Sa - Sl + bl, \\ &-2R + Z + x, -2S + W + y, x^2b - xbl + y^2b - ya^2 + yal - yb^2, \\ &XWt + Xty - YZt + Ytx - Zty - Wtx - 1); \end{aligned}$$

but we obtain

$$\mathsf{NormalForm}(1, \mathsf{I}) = 1.$$

Surprisingly, we must conclude that the theorem is not geometrically true, at least in the way it has been algebraically translated.

1.2. NON-DEGENERACY CONDITIONS

Algebraic varieties, namely, zero sets of algebraic equations, are 'small' or 'thin' subsets of the affine space they lie in. In fact, they are considered small, since they do not include any hypercube or hypersphere of that affine space: a planar curve has no width at any point, a surface in space is never thick, and so on. On the other hand, algebraic varieties are always closed sets in the Euclidean topology of the real or complex affine space. But there is a different topology that can be constructed on any given algebraic variety: the Zariski topology, having as closed sets precisely those algebraic varieties contained in the given one. Thus, open Zariski sets are to be considered as 'large', being the complement of an algebraic variety.

These technical remarks allow us to analyze some statements that fail to be geometrically true. Concretely, statements that do not hold in all cases but that are true if a few geometric instances of the hypotheses are ruled out‡ (that is, if they fail only on a 'small' set of points of the hypothesis variety). Technically speaking, we are considering now that H is not a subset of T, but that a nonempty Zariski open subset $W \subseteq H$ is contained in T. The important news is that we can still detect such statements, performing a small modification of the above procedure. Namely, when the theorem is not geometrically true, we should compute a basis of the ideal

$$\mathsf{J} = ((h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t]) \cap K[x_1, \ldots, x_n].^{‡‡}$$

It is immediate from a precedent observation that for every $g(x_1, \ldots, x_n)$ in this ideal, and any $\bar{x} = (x_1, \ldots, x_n) \in L^n$, it holds that

* Freely distributed at (cocoa@dima.unige.it).

** The Normal Form of an element with respect to some ideal happens to be zero iff the element belongs to the ideal; see [4].

‡ See the discussion at the end of 1.3 for a more precise description.

‡‡ This operation of finding the polynomials in an ideal that include just some specific variables is called the *elimination* of the remaining variables.

$$h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge g(\bar{x}) \neq 0 \Rightarrow k(\bar{x}) = 0.$$

Therefore, elements of this ideal J provide some complementary hypotheses of a negative kind ($g \neq 0$) such that, adding them to the given hypotheses, the thesis now follows. Of course, to avoid proving theorems that hold because of contradictory hypotheses, it should be required as well that the new set of conditions $h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge g(\bar{x}) \neq 0$ is nonempty (i.e., that $H \not\subset \{g = 0\}$). For reasons that will become apparent soon, these extra hypotheses $g \in$ J are often called *non-degeneracy conditions** and the ideal J is referred to as the *non-degeneracy ideal.* If $H \subset \{g = 0\}$, the non-degeneracy condition $g \neq 0$ is called *trivial.* Thus we should always search for nontrivial conditions.

This extended procedure can be algorithmically carried on computing (as in [7]), by elimination, a basis of J, say, $(g_1, \ldots, g_s)$, for instance, through the command Elim from CoCoA. Next we should look for trivial generators in this basis, and this can be done obtaining, for each i, the normal form of 1 in the ideal $(h_1, \ldots, h_p, g_i t - 1)$ and determining – as above – whether H is contained in the zero-set $\{g_i = 0\}$. If $H \subset \bigcap_{i=1}^{s} \{g_i = 0\}$, then all conditions in the found basis are trivial; moreover, it is easy to see that this implies that all conditions in the ideal of non-degeneracy are trivial as well.

If this is not the case, there will be points $\bar{x} = (x_1, \ldots x_n) \in L^n$ such that

$$h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge (g_1(\bar{x}) \neq 0 \vee \cdots \vee g_s(\bar{x}) \neq 0)$$

and for all such $\bar{x} = (x_1, \ldots, x_n) \in L^n$,

$$h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge (g_1(\bar{x}) \neq 0 \vee \cdots \vee g_s(\bar{x}) \neq 0) \Rightarrow k(\bar{x}) = 0$$

so that the theorem now holds under any of these non-degeneracy conditions.

Let us see how this technique applies to our previous example. Performing the elimination of the slack variable t introduced in the last line of the definition of I, one gets a collection of over thirty polynomials, say, J = Elim(t, I), in the remaining variables $\{LMRSXYZWxyabl\}$, too large to reproduce here. Some of them yield trivial conditions, for instance, all the generators of the hypotheses ideal (all the generators of I excluding the last one), which are obviously contained in J. But some others polynomials from the elimination basis, for instance, $a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2$ give nontrivial conditions:

```
NormalForm(1, Ideal(La + Mb - xa - yb, -Lb + Ma, -2L + X + x,
              -2M + Y + y, Ra - Rl + Sb - xa + xl - yb,
              -Rb + Sa - Sl + bl, -2R + Z + x, -2S + W + y,
              x^2b - xbl + y^2b - ya^2 + yal - yb^2)
              (a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2)t - 1);
     1;
```

* We are, for the sake of clarity, omitting in this subsection some technical precisions regarding non-degeneracy, in particular, the fact that it should be defined only in terms of independent variables; but we want to postpone the introduction of this elusive concept until the end of 1.3 and Section 2.

Therefore the theorem should hold under the condition $a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2 \neq 0$. We check this adding to the ideal I the polynomial $(a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2)h - 1$, where h is another – different – slack variable. As expected, now the normal form of 1 is 0.

$$\begin{aligned}\text{NormalForm}(1,\ \text{Ideal}(&La + Mb - xa - yb, -Lb + Ma, -2L + X + x,\\ &-2M + Y + y, Ra - Rl + Sb - xa + xl - yb,\\ &-Rb + Sa - Sl + bl, -2R + Z + x, -2S + W + y,\\ &x^2b - xbl + y^2b - ya^2 + yal - yb^2)\\ &(a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2)t - 1,\\ &XWt + Xty - YZt + Ytx - Zty - Wtx - 1));\end{aligned}$$

0;

What does this non-degeneracy condition mean geometrically? It factors as $((a-l)^2 + b^2)(a^2 + b^2) \neq 0$, that is, $\{(a-l)^2 + b^2 \neq 0\} \wedge \{a^2 + b^2 \neq 0\}$. So the first clause in this conjunction can be interpreted, over the reals, as $\{a \neq l\} \vee \{b \neq 0\}$, and the second one means $\{a \neq 0\} \vee \{b \neq 0\}$. Condition $\{b \neq 0\}$ implies the given triangle does not collapse to a certain line; the other two conditions imply the triangle is not straight and could be considered as degeneracies associated to the algebraic formulation. We warn the reader about imposing, instead of the found condition, any of these 'real' counterparts: they do not yield a 0 normal form. Although they seem more natural in the geometric context, they do not translate well the algebraic behavior of the geometric statement. These real-field interpretations should be 'for your eyes only'. A similar analysis can be carried on with other conditions arising from this elimination computation.

In summary, in the case that there are nontrivial non-degeneracy conditions and the ideal of non-degeneracy has a basis $(g_1, \ldots, g_s)$, with, say, $(g_1, \ldots, g_r)$ nontrivial, the given theorem holds over the nonempty open set of $H \wedge (g_1(\bar{x}) \neq 0 \vee \cdots \vee g_r(\bar{x}) \neq 0)$ (which happens to be the largest open one in H where the theorem holds).

1.3. PROVING PSEUDO-THEOREMS

Unfortunately, things are a little bit more complicated than the previous example shows: in fact, if we apply mechanically the procedures we have developed in the preceding sections, it can happen that we end up labeling as truthful a statement that happens to hold only over cases that are, intuitively, degenerate. Such a phenomenon is what we informally name as proving 'pseudo-theorems'. Let us describe what we mean through a suitable modification of that example's formulation.

EXAMPLE 2. *This time we will consider as data the given triangle* O$(0, 0)$, A$(l, 0)$, B(a, b) *and an arbitrary point* x(x, y). *We also consider the symmetrical images of such point with respect to the three sides of the triangle,* x′$(x, -y)$, x″(X, Y), x‴(Z, W), *but we forget to input, as one of the hypotheses, that the point* x(x, y) *is* *in the circumcircle of the triangle. Still, we are claiming that the three symmetrical images of* x *are aligned, that is,* $XW - Zy + xY + yX - YZ - xW = 0$.

First, as expected, we find out that the theorem is not geometrically true (and we will omit this part of the procedure); then, in order to find non-degeneracy conditions, we denote by I′ the ideal

$$\begin{aligned} \mathsf{I}' = \ & \mathsf{Ideal}(a(L-x)+b(M-y), -bL+aM, X-x-2(L-x), \\ & Y-y-2(M-y), (a-l)(R-x)+b(S-y), \\ & -b(R-l)+(a-l)S, Z-x-2(R-x), W-y-2(S-y), \\ & (XW-Zy+xY+yX-YZ-xW)t-1) \end{aligned}$$

and proceed eliminating t. Let $\mathsf{J}' = \mathsf{Elim}(t, \mathsf{I}')$ be the elimination ideal. Then we obtain

$$\begin{aligned} \mathsf{J}' = \ & \mathsf{Ideal}(Xa+Yb-xa-yb, Xb-Ya+xb-ya, L-1/2X-1/2x, \\ & M-1/2Y-1/2y, Za-Zl+Wb-xa+xl-yb, \\ & Zb-Wa+Wl+xb-ya+yl-2bl, R-1/2Z-1/2x, \\ & S-1/2W-1/2y, Z^2b-2Zbl+W^2b-x^2b+2xbl-y^2b, \\ & 1/2XZl-XWb-1/2Xxl+YZb+1/2YWl+1/2Yyl- \\ & -Ybl-1/2Zxl+1/2Wyl+1/2x^2l+1/2y^2l-ybl, \\ & XZb+1/2XWl+1/2Xyl-Xbl-1/2YZl+YWb+ \\ & +1/2Yxl-1/2Zyl-1/2Wxl-x^2b+xbl-y^2b, \\ & 1/2XWl+Xxb+1/2Xyl-Xbl+1/2YZl-1/2Yxl+Yyb- \\ & -Zxb+1/2Zyl-1/2Wxl-Wyb-xyl+xbl, \\ & 1/2XZl-1/2Xxl+Xyb-1/2YWl-Yxb-1/2Yyl+Ybl- \\ & -1/2Zxl-Zyb+Wxb-1/2Wyl+1/2x^2l-1/2y^2l+ybl, \\ & X^2b+Y^2b-x^2b-y^2b, Z^2l-2Zl^2+W^2l-x^2l+2xl^2-y^2l, \\ & X^2l+Y^2l-x^2l-y^2l); \end{aligned}$$

After some computations, we can check that all but the last two polynomials yield trivial conditions (in the sense of 1.2). That neither $Z^2l-2Zl^2+W^2l-x^2l+2xl^2-y^2l$ nor $X^2l+Y^2l-x^2l-y^2l$ is trivial means that each of the corresponding zero sets $Z^2l-2Zl^2+W^2l-x^2l+2xl^2-y^2l=0, \quad X^2l+Y^2l-x^2l-y^2l=0$ does not vanish identically over the hypothesis variety.

Factorizing the equations, we see that, geometrically, $Z^2l-2Zl^2+W^2l-x^2l+2xl^2-y^2l=0$ expresses the disjunction

$$\{l=0\} \vee \{d((Z,W),(l,0)) = d((x,y),(l,0))\}$$

while $X^2l+Y^2l-x^2l-y^2l=0$ means

$$\{l=0\} \vee \{d((X,Y),(0,0)) = d((x,y),(0,0))\}.$$

Thus there is a Zariski-open set of points in the hypothesis variety where both $\{l \neq 0\} \wedge \{d((Z,W),(l,0)) \neq d((x,y),(l,0))\}$ (and another open set where $\{l \neq 0\} \wedge$

$\{d((X, Y), (0, 0)) = d((x, y), (0, 0))\}$). And over each one of these open sets the given statement holds, because the polynomials defining the open sets belong to the non-degeneracy ideal. Therefore we must conclude that the three symmetrical images x', x'', x''' of *any point* x, with respect to the sides of a triangle, are aligned if $\{l \neq 0\}$, that is, the triangle does not collapse to a certain line, *and* if $d(x''', (l, 0)) \neq d(x, (l, 0))$ or $d(x'', (0, 0)) \neq d(x, (0, 0))$.

Thus we have discovered a property that holds over some nonempty open sets of the hypothesis variety, but it is intuitively obvious – if we draw some pencil and paper sketch of this geometric construction – that something goes wrong here! First of all we notice that $(l, 0)$ (resp. $(0, 0)$) is in the symmetry axis between (Z, W) and (x, y) (resp. (X, Y) and (x, y)). And symmetrical points lie at equal distance from any point in the symmetry axis. Therefore, over 'normal' triangles one would have expected that both equalities $d((Z, W), (l, 0)) = d((x, y), (l, 0))$ and $d((X, Y), (0, 0)) = d((x, y), (0, 0))$ should always hold. Since the open sets where we have discovered our theorem to hold do not verify such equalities, we must conclude that these open sets include only intuitively 'abnormal' triangles.

Aiming to understand better what is happening, let us exhibit what kind of triangles lie in these open sets. Thus we consider the ideal generated by the original hypotheses and, say, one of these non-degeneracy conditions (expressed as $X^2l + Y^2l - x^2l - y^2l \neq 0$), and we eliminate all the variables but those that seem *independently* given in our construction – in the intuitive sense that their values can be arbitrarily determined and also that they finitely determine the values of the remaining ones – namely, $\{xyabl\}$. Then we will get polynomials that, being combinations of the hypotheses and of the non-degeneracy conditions, should vanish over all points in the corresponding open set. On the other hand, these polynomials will involve only independent variables, and independency means that no nontrivial algebraic combination of these variables should vanish on open sets of intuitively non-degenerate instances of the hypothesis variety. Both considerations imply that, if the open sets we are analyzing do include some reasonable cases of our geometric construction, the elimination must yield the zero ideal. But let us see what we obtain here:

$$\begin{aligned}&\mathsf{Elim}(t..W, \mathsf{Ideal}((X^2 + Y^2 - x^2 - y^2)lt - 1, a(L - x) + b(M - y),\\ &\quad - bL + aM, X - x - 2(L - x), Y - y - 2(M - y),\\ &\quad (a - l)(R - x) + b(S - y), -b(R - l) + (a - l)S,\\ &\quad Z - x - 2(R - x), W - y - 2(S - y)));\\ &\mathsf{Ideal}(b, a);\end{aligned}$$

Likewise, for the other non-degeneracy condition we get

$$\mathsf{Ideal}(b, a - l).$$

Thus we see that all triangles verifying any of the two non-degeneracy conditions necessarily verify also $b = 0$, so they must collapse to a line (remark that this happens even imposing that $\{l \neq 0\}$). We can conclude that the open sets where we have discovered that the stated thesis (the three symmetrical points are aligned) does hold are also subsets of the hypothesis variety where all triangles collapse to lines.*

We could say that we have proved a 'pseudo-theorem': the statement, as it has been algebraically translated, is true in some open set, but of geometrically degenerate cases. On the other hand, the example in 1.2 was also a statement holding on an open set, but of geometrically non-degenerate cases. Thus, the difference between the two cases depends on the idea of geometric degeneracy, perhaps clear for human intuition, but difficult to formulate in objective terms. For instance, we could use human cooperation during the process of translating into algebraic formulae the given geometric hypotheses, so that the *independent* variables are explicitly declared before starting the automatic proving procedure. By definition, no nonzero polynomial in the independent variables should vanish on any open set of geometrically non-degenerate cases. In this framework we could declare as pseudo-theorems** those holding just over open subsets of the hypothesis variety where the independent variables verify some nontrivial algebraic relation among them, such as $b = 0$ in the above example. Since this procedure to rule out pseudo-theorems is well documented in the bibliography, we present here a short summary of the situation, and we refer the reader to any of the references (in particular to [1]) for complete details.

Briefly, the intuitive concept of independent variables is technically formulated as the concept of algebraically independent variables modulo the hypothesis ideal, that is, variables $\{x_1, \ldots, x_d\}$ such that $(0) = ((h_1, \ldots, h_p)K[x_1, \ldots, x_n,]) \cap K[x_1, \ldots, x_d]$. In algebraic-geometry terms, independent variables are those verifying that no nonzero polynomial in these variables vanishes over some irreducible component of the hypothesis variety (equivalently, over any nonempty open subset of such component). There are, in general, several possible maximal sets of independent variables, even with maximal cardinality. After a meaningful set of independent variables, say, $\bar{x}' = (x_1, \ldots, x_d)$, has been identified – through hu-

* In fact, we see that the originally given hypothesis variety has many other peculiar features: for instance, it has dimension higher than expected: it 'should' have been of dimension five, since there are just five independent variables $\{xyabl\}$ that geometrically control the construction; but, actually, it has dimension six, as we can check, again, with CoCoA (P stands for the name of the current ring with all the involved variables $\{LMRSXYZWxyabl\}$:

$$\begin{aligned}\text{dim(P/Ideal } &(a(L-x)+b(M-y), -bL+aM, X-x-2(L-x),\\ &Y-y-2(M-y), (a-l)(R-x)+b(S-y), -b(R-l)+(a-l)S,\\ &Z-x-2(R-x), W-y-2(S-y)))\end{aligned}$$

6:

Moreover, we can verify that this dimension diminishes by one if b is required to be nonzero, in other words, if we restrict to the set of points in the hypothesis variety that correspond to 'normal' triangles.

** See the formal definition of generally false theorems in the next section.

man intervention – among the complete set of variables $(x_1, \ldots, x_n)$, the ideal of non-degeneracy should be redefined as the elimination ideal

$$\mathrm{J} = ((h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t]) \cap K[x_1, \ldots, x_d].$$

Accordingly, the irreducible components of the hypothesis variety where $\{x_1, \ldots, x_d\}$ remain independent are now labeled as *non-degenerate components*.

This time the nonzero elements of J always yield nontrivial conditions, since they can belong neither to some components of the hypothesis ideal and nor to the ideal itself. Again, it can be shown that for every nonzero element $g(\bar{x}') \in \mathrm{J}$, and for all $\bar{x} = (x_1, \ldots, x_n) \in L^n$,

$$h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge g(\bar{x}') \neq 0 \Rightarrow k(\bar{x}) = 0$$

but this time we are sure that the set

$$h_1(\bar{x}) = 0 \wedge \cdots \wedge h_p(\bar{x}) = 0 \wedge g(\bar{x}') \neq 0$$

includes all non-degenerate components: in fact, over such components g cannot be identically zero, so k must vanish over a nonempty open set of the component, and therefore, by its irreducibility, over the whole component.

Thus, the existence of nonzero element in

$$((h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t]) \cap K[x_1, \ldots, x_d]$$

is equivalent to proving the given thesis holds over all non-degenerate irreducible components, and this set surely includes an open set of cases where nonzero polynomials in the independent variables cannot vanish identically. This is what happened in Example 1, since we have found a non-degeneracy condition in Subsection 1.2, in terms of the independent variables. It is said, then, that the theorem is *generally true*, that is, true under some negative subsidiary condition on the independent variables. On the other hand, it is easy to see that there is no such non-degeneracy condition for Example 2, and therefore we can say this is an example of a *not generally true* theorem. Finally, we must notice that there are statements (see Example 4) that hold over some open set of non-degenerate cases, but not over all non-degenerate irreducible components. They should not be considered as pseudo-theorems, but still they will be simply called *not generally true*. Making a finer filter to separate all three cases requires some more advanced techniques that will be briefly discussed later on the next section.

2. Automatic Discovery

2.1. THE METHOD

From the point of view of exploring open geometric situations, our interpretation of the above (more or less classical) results is the following. Assume hypotheses and theses are given in some totally arbitrary way, then the computer will output answers of three possible kinds: (1) the conjectured theorem is true as stated; (2) it is (generally) true; (3) could be that it is not generally true, even when considering any complementary non-degeneracy hypotheses, if they are all trivial (i.e., if the ideal of non-degeneracy conditions is zero).

Therefore, with the outlined method we will be able to deal only with those theorems that could be previously 'nearly' guessed, since – with the terminology introduced at the beginning of 1.2 – the thesis should hold over a 'large' (open) subset of the hypothesis variety for the method to output some helpful information. So we must, in practice, know in advance that the theorem is 'almost' true and the method just confirms/slightly corrects/refutes the correctness of the initial guess. Therefore, up to now the method has been, properly speaking, one of automatic 'proving'. If our initial guess is wrong, the preceding approach does not provide any useful hint about how to proceed any further. On the other hand, we could say that automatic theorem 'discovery' should deal, in most cases, with detecting just a 'thin' subset of the hypothesis variety, not a large one. In fact, if the thesis does not hold over any large set of the hypothesis variety, it could still happen that the thesis vanishes over a smaller subset of the hypothesis variety; or it could be that it is not true at all over any point. But, so far, we have not provided any means to identify such truth-subsets in this particularly interesting situation when we have not made a correct guess of a geometric theorem.

The method for automatic discovery of elementary geometry theorems that we propose here is quite straightforward. Let us assume that a given thesis k does not hold over any open set (described by polynomials in the geometrically independent variables $\bar{x}' = (x_1, \ldots, x_d)$) of the hypothesis variety $\{h_1 = 0, \ldots, h_p = 0\}$, that is, that the theorem is not generally true, so that

$$(0) = (h_1, \ldots, h_p, kt - 1)K[x_1, \ldots, x_n, t] \cap K[x_1, \ldots, x_d].$$

Then, we simply start by adding the thesis to the collection of hypotheses. Of course, since $T \wedge H \Rightarrow T$ is always true, it is obvious that whatever the thesis might be, it will now follow from the enlarged set of hypotheses. In geometric terms, the intersection of the hypothesis variety and the thesis hypersurface is always included in the latter variety. The key step toward discovery is to rephrase the conjunction of hypotheses *and* thesis in terms of geometrically meaningful variables, namely, the independent ones. Thus we must eliminate the nonindependent variables from the new ideal of (*hypotheses*, *thesis*). The vanishing of every element h' in this elimination ideal (*hypotheses*, *thesis*)$\cap K$[*independent variables*] is obviously a necessary condition for the theorem to hold, since it is a combination of (*hypotheses*, *thesis*).

If the theorem holds over some instance, then both hypotheses and thesis must vanish on it, and the same should happen for the elements of the elimination ideal.

Then we have to distinguish two possible cases.

- First, that $(h'_1, \ldots, h'_r) = (h_1, \ldots, h_p, k) \cap K[x_1, \ldots, x_d] \neq (0)$.

  It is easy to show* that this happens if and only if k vanishes on none of the non-degenerate components of the (old) hypothesis variety. This situation has already been labeled in the literature ([1]) as the *generally false* case. In fact it means that the theorem does not hold over any open set of geometrically non-degenerate cases: it can only hold over open sets of geometrically degenerate instances.**

  In this case, moreover, the zero set of $(h'_1, \ldots, h'_r)$ in the former hypothesis variety gives a proper closed set (since it includes points only where the variables in $\bar{x}'$ are dependent). So we have now a strictly smaller hypothesis variety given by $(h_1, \ldots, h_p, h'_1, \ldots, h'_r)$ that it is more likely to be contained in the thesis variety. Next we should restart again the automatic theorem-proving method, identifying over this new hypothesis variety a subset $\{x_1, \ldots, x_m\}$, where $m \leq d$, of independent variables, testing whether the thesis is generally true over such variety and finding, if that is the case, an open set of non-degenerate conditions, and so on.

  Note that over the new hypothesis variety the given thesis is always not generally false,‡ so unless it is generally true over this variety, we will be lead into the next situation.
- Second, that $(0) = (h_1, \ldots, h_p, k) \cap K[x_1, \ldots, x_d]$ (the not generally false case). Then we know that the thesis holds over some non-degenerate component, but also that it does not hold over some other non-degenerate component (for we are assuming the given statement is not generally true). In such situation we do not know how to proceed further on without decomposing the variety into irreducible components, and we do not consider feasible to get into such computational problem at this moment.

Summarizing, the method could be outlined as follows:

Input: Hypotheses as algebraic equations $\{h_1 = 0, \ldots, h_p = 0\}$, and a conclusion $k = 0$, also an algebraic equation, all over variables $x_1, \ldots, x_n$.

---

* Take $0 \neq g \in (h_1, \ldots, h_p, k) \cap K[x_1, \ldots, x_d]$. Then $g \neq 0$ over the non-degenerate components and, since g is a combination of k, h_i's and all h_i's are zero over every component, it follows that $k \neq 0$ over the non-degenerate components. Conversely, if k vanishes on none of the non-degenerate components, we construct a polynomial $g \in K[\bar{x}']$ as follows: take, for each of the remaining components, a polynomial $g_i(\bar{x}')$ that vanishes over it, and let g be the product of all g_i's. Then $\{h_1 = 0, \ldots, h_p = 0, k = 0\} \subset \{g = 0\}$ and thus $g \in \sqrt{(h_1, \ldots, h_p, k)} \cap K[\bar{x}']$, which implies $(h_1, \ldots, h_p, k) \cap K[\bar{x}'] \neq 0$.

** Such as the example in 1.3.

‡ Obviously, $(0) = (h_1, \ldots, h_p, h'_1, \ldots, h'_r, k) \cap K[x_1, \ldots, x_m] = (h'_1, \ldots, h'_r) \cap K[x_1, \ldots, x_m]$, by the independence of the variables $\{x_1, \ldots, x_m\}$.

(1) Check whether the hypotheses are contradictory by computing
NormalForm$(1, \text{Ideal}(h_1, \ldots, h_p))$.

If the answer is 0, *output* a warning message. Otherwise, find a maximal set of geometrically meaningful, independent variables $\{x_1, \ldots, x_d\}$ over the hypothesis variety.

(2) Check whether the theorem is geometrically true, computing
NormalForm$(1, \text{Ideal}(h_1, \ldots, h_p kt - 1))$.

If this normal form is zero, *output*: theorem geometrically true.

(3) If not geometrically true, check whether the theorem is generally true, computing
Elim $(t \ldots x_{d+1}, \text{Ideal}(h_1, \ldots, h_p, kt - 1))$.

If the answer is a set of nonzero polynomials $\{g_1, \ldots, g_s\}$ in the variables $\{x_1, \ldots, x_d\}$, *output*: theorem generally true under non-degeneracy conditions $g_1 \neq 0 \vee \cdots \vee g_s \neq 0$.

(4) Else, the theorem is not generally true and we look for missing hypotheses, computing
Elim $(x_n \ldots x_{d+1}, \text{Ideal}(h_1, \ldots, h_p, k))$.

If the answer is zero, *output*: theorem not generally true and not generally false (decomposition required!).

(5) If the answer is a set of nonzero polynomials $\{h'_1, \ldots, h'_r\}$, go to Step 1, considering now the set of hypotheses $\{h_1, \ldots, h_p, h'_1, \ldots, h'_r\}$.

We remark that Step 5 can be improved in some cases: for instance, if only one nonzero missing hypothesis h' is obtained and it can be factorized, then instead of going to Step 1 adding h', we could add just anyone of its factors (it yields a smaller hypothesis variety).

As can be deduced from the above description, the method is not complete. This is shown in the example of Subsection 2.3 (where we have arrived at a dead end after one successful step of the outlined method).

2.2. A SUCCESSFUL CASE: EXAMPLE 2 REVISITED

EXAMPLE 3. Same hypotheses and thesis of Example 2 of 1.3.

After checking in 1.3 that the theorem is not generally true, let us call I″ the ideal generated by the set of hypotheses and by the thesis (alignment of the three symmetrical images of x):

$$\begin{aligned} \text{I}'' = \text{Ideal}(&a(L-x)+b(M-y), -bL+aM, X-x-2(L-x),\\ &Y-y-2(M-y), (a-l)(R-x)+b(S-y),\\ &-b(R-l)+(a-l)S, Z-x-2(R-x),\\ &W-y-2(S-y), XW-Zy+xY+yX-YZ-xW). \end{aligned}$$

Let $J'' = \mathrm{Elim}(L..W, I'')$ be the corresponding elimination ideal that removes all variables except the independent ones $\{xyabl\}$:

$$J'' = \mathrm{Ideal}(x^2b^3l - xb^3l^2 + y^2b^3l - ya^2b^2l + yab^2l^2 - yb^4l).$$

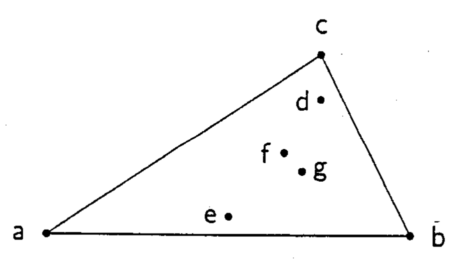
Thus we are in a generally false case. Equating to zero the only generator of this elimination ideal, we express, therefore, a necessary condition for the alignment of the symmetrical points of (x, y), namely, $h' = x^2b^3l - xb^3l^2 + y^2b^3l - ya^2b^2l + yab^2l^2 - yb^4l = 0$. It can be interpreted as $\{b^2l = 0\}^* \vee \{(x, y)$ lies in the circumcircle of the given triangle$\}$. We should continue then with the automatic proving procedure, checking if the thesis $XW + Xy - YZ + Yx - Zy - Wx = 0$ is generally true with the given set of hypotheses plus $h' = 0$ and considering as new set of independent variables $\{yabl\}$, since $\{xyabl\}$ are already related by $h' = 0$.

$$\begin{aligned}\mathrm{Elim}(t..x, \mathrm{Ideal}(&La + Mb - xa - yb, -Lb + Ma, -2L + X + x,\\ &-2M + Y + y, Ra - Rl + Sb - xa + xl - yb,\\ &-Rb + Sa - Sl + bl, -2R + Z + x, -2S + W + y,\\ &x^2b^3l - xb^3l^2 + y^2b^3l - ya^2b^2l + yab^2l^2 - yb^4l,\\ &(XW + Xy - YZ + Yx - Zy - Wx)t - 1))\\ \mathrm{Ideal}(a^4 - 2a^3l &+ 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2);\end{aligned}$$

We conclude that the theorem, with the new hypothesis, is generally true under the condition $a^4 - 2a^3l + 2a^2b^2 + a^2l^2 - 2ab^2l + b^4 + b^2l^2 \neq 0$, which is the same that appeared in 1.3 for Example 1. Therefore, over (non-degenerate) triangles that verify $\{b^2l \neq 0\}$ we have discovered a new hypothesis for our thesis to hold: namely, that the given point (x, y) must lie in the outer circle of the triangle.

2.3. A CASE OF FAILURE

EXAMPLE 4. *In a triangle, the orthocenter (intersection of heights), the centroid (intersection of medians), the circumcenter and the incenter (center of the excircle or outer circle, resp. incircle or inner circle, of the triangle) lie on a line.*



Let us consider the triangle of vertices a$(-1, 0)$, b$(1, 0)$, c(a, b). Let d(p, q) be the orthocenter, e(u, v) the circumcenter, and f(l, r) the centroid, respectively.

* From the intuitive point of view, a case of degeneracy.

Denote by o the origin $(0, 0)$. Now we state the algebraic conditions verified by such points:

$$\begin{array}{lll} \textit{Hypotheses}: & \overline{\mathsf{ab}} \perp \overline{\mathsf{cd}}: & p - a = 0 \\ & \overline{\mathsf{ac}} \perp \overline{\mathsf{bd}}: & (p-1)(a+1) + qb = 0 \\ & \overline{\mathsf{ae}} = \overline{\mathsf{ce}}: & (u+1)^2 + v^2 - (u-a)^2 - (v-b)^2 = 0 \\ & \overline{\mathsf{ae}} = \overline{\mathsf{be}}: & (u+1)^2 - (u-1)^2 = 0 \\ & \mathsf{f} \in \overline{\mathsf{co}}: & ar - bl = 0 \\ & \mathsf{f} \in \overline{\mathsf{b(a+c)/2}} & (l-1)b - r(a-3) = 0 \end{array}$$

Moreover, the incenter $\mathsf{g}(s, w)$ verifies that it is the center of a circle of radius w, which is also tangent to the sides of the triangle. We use CoCoA to find determining equations for this point by eliminating variables $\{x, y\}$ in the two systems each consisting of the equation of the circle of center (s, w) and radius w, $(x-s)^2+(y-w)^2 - w^2 = 0$, and of the equation expressing the perpendicularity from a radius of this circle to the side $\overline{\mathsf{ac}}$ (resp. $\overline{\mathsf{bc}}$): $b(x+1) - (a+1)y$ (resp. $b(x-1) - (a-1)y$).

Let us use the set $\{txyswpquvlrab\}$ as variables for all the computations concerning the example, remarking that here $\{a, b\}$ are the only independent variables (taking t as an auxiliary variable in some computations). Then the determining equations for the incenter can be obtained as follows:

$$\begin{array}{l} \mathsf{Elim}(t..y, \mathsf{Ideal}((x-s)^2 + (y-w)^2 - w^2, b(x+1) - (a+1)y, \\ \qquad (x-s)(a+1) + (y-w)b)); \\ \mathsf{Ideal}(swab - 1/2s^2b^2 + 1/2w^2b^2 + swb + wab - sb^2 + wb - 1/2b^2); \end{array}$$

$$\begin{array}{l} \mathsf{Elim}(t..y, \mathsf{Ideal}((x-s)^2 + (y-w)^2 - w^2, b(x-1) - (a-1)y, \\ \qquad (x-s)(a-1) + (y-w)b)); \\ \mathsf{Ideal}(swab - 1/2s^2b^2 + 1/2w^2b^2 - swb - wab + sb^2 + wb - 1/2b^2); \end{array}$$

Therefore, we can take the two output polynomials as the hypothesis equations that simultaneously determine g. As it is well known, the centroid, orthocenter, and circumcenter are aligned. Let us focus on the following statement: the incenter $\mathsf{g}(s, w)$, the circumcenter $\mathsf{e}(u, v)$, and the centroid $\mathsf{f}(l, r)$ lie in a line, namely, they verify $(sv + ur + lw - rs - wu - lv) = 0$. After noticing that this theorem is not generally true, we apply our technique of adding the thesis itself to the collection of hypotheses and eliminating

$$\begin{array}{l} \mathsf{Elim}(t..r, \mathsf{Ideal}((u-1)^2 + v^2 - (u-a)^2 - (v-b)^2, (u+1)^2 - (u-1)^2, \\ \qquad ar - bl, (l-1)b - r(a-3), \\ \qquad swab - 1/2s^2b^2 + 1/2w^2b^2 + swb + wab - sb^2 + wb - 1/2b^2, \\ \qquad swab - 1/2s^2b^2 + 1/2w^2b^2 - swb - wab + sb^2 + wb - 1/2b^2, \\ \qquad sv + ur + lw - rs - wu - lv)); \end{array}$$

$$\begin{array}{l} \mathsf{Ideal}(a^5b + 2a^3b^3 + ab^5 - 10a^3b - 6ab^3 + 9ab, \\ a^7 - 3a^3b^4 - 2ab^6 - 11a^5 + 12a^3b^2 + 11ab^4 + 19a^3 - 12ab^2 - 9a); \end{array}$$

Thus, a necessary condition for the validity of our theorem is that these last two polynomials vanish simultaneously. In the previous example it was the case that the discovered necessary conditions were also sufficient, but not in this example: as shown below, we cannot find non-degeneracy hypothesis for the alignment thesis to hold over triangles verifying $\{a^5b + 2a^3b^3 + ab^5 - 10a^3b - 6ab^3 + 9ab = 0\} \wedge \{a^7 - 3a^3b^4 - 2ab^6 - 11a^5 + 12a^3b^2 + 11ab^4 + 19a^3 - 12ab^2 - 9a = 0\}$.* Note that now only, say, $\{b\}$, is to be considered as independent variable, since $\{a, b\}$ are linked by one equation (but it yields the same result considering a instead).

$$
\begin{aligned}
&\mathsf{Elim}(t..a, \mathsf{Ideal}((u-1)^2 + v^2 - (u-a)^2 - (v-b)^2,\\
&\quad (u+1)^2 - (u-1)^2, ar - bl, (l-1)b - r(a-3),\\
&\quad swab - 1/2s^2b^2 + 1/2w^2b^2 + swb + wab - sb^2 + wb - 1/2b^2,\\
&\quad swab - 1/2s^2b^2 + 1/2w^2b^2 - swb - wab + sb^2 + wb - 1/2b^2,\\
&\quad a^5b + 2a^3b^3 + ab^5 - 10a^3b - 6ab^3 + 9ab,\\
&\quad a^7 - 3a^3b^4 - 2ab^6 - 11a^5 + 12a^3b^2 + 11ab^4 +\\
&\quad + 19a^3 - 12ab^2 - 9a,\\
&\quad (sv + ur + lw - rs - wu - lv)t - 1));\\
&\mathsf{Ideal}(0);
\end{aligned}
$$

Therefore it turns out that the alignment of incenter, centroid, orthocenter and circumcenter on isosceles triangles cannot be proved, with the proposed method, to be generally true under the new, discovered hypotheses. As remarked in 2.1, it is not either generally false. Thus we are stuck, and our method is shown to be imcomplete.

2.4. DISCOVERY THROUGH FAILURE

Nevertheless, the reader probably believes that the claimed theorem holds over isosceles triangles, after he/she tests a few examples. We have arrived at a paradoxical situation that provides, simultaneously, one example of the limitations of the method (since we have here a statement that cannot be conclusively analyzed) and, on the other hand, one example of its success (we have found an 'intuitively' true theorem, holding over some specific kind of triangles discovered by this technique)!

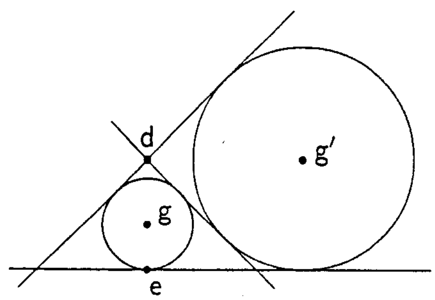
---

* In fact, the two extra hypotheses can be factorized as follows:

$$
\begin{aligned}
a^5b &+ 2a^3b^3 + ab^5 - 10a^3b - 6ab^3 + 9ab\\
&= ab(a^2 + 2a - 3 + b^2)(a^2 - 2a - 3 + b^2),\\
a^7 &- 3a^3b^4 - 2ab^6 - 11a^5 + 12a^3b^2 + 11ab^4 + 19a^3 - 12ab^2 - 9a\\
&= a(a^2 + 2a - 3 + b^2)(a^2 - 2a - 3 + b^2)(a^2 - 1 - 2b^2).
\end{aligned}
$$

We observe that each of the factors of the GCD of these two polynomials express the condition: *the triangle must be isosceles.*

The reason behind such paradox is that our algebraic description of incenter point does not actually determine the true incenter we are intuitively thinking of, but, instead, a collection of four points (the centers of the four tri-tangent circles to a triangle), since the incenter was given as an intersection of two degree-two equations in s, w. Obviously, it is not possible that this four points are simultaneously aligned with the other given points (centroid, circumcenter, orthocenter), since they are never aligned themselves (see the figure, where we have displayed just two incenters and the points d and e, which are aligned with only some of the incenters)!



The algebraic-geometry counterpart of this fact is that we have several components for the different incenters. When we project from the $xyswpquvlrab$ space onto the $a - b$ plane the original hypothesis variety intersected with the equation expressing the thesis, we get three curves: $a = 0$, $(a^2 + 2a - 3 + b^2) = 0$, $(a^2 - 2a - 3 + b^2) = 0$, each one prescribing different pairs of sides of the triangle to be equal. For each of these equality cases, two incenters become aligned with the orthocenter, centroid and circumcenter. But when we consider the cylinder over anyone of these curves, it comes out that it intersects some of the components of the hypothesis variety (but not that corresponding to the 'internal' incenter) in points (associated to positions of the other incenters that are not aligned) that are not in the variety intersected with the thesis; therefore some components of the intersection of this cylinder with the variety are not contained in the thesis hypersurface, but one is. When we work over the component of the internal incenter, the lifting of the $a - b$ projection does not include any extraneous point, and thus the thesis surely holds over this component, as intuitively expected. We remark that when we consider some other component, then only for the lifting of some concrete curve (such as $a^2 - 2a - 3 + b^2 = 0$) of the projection does not include points not in the thesis, but the other curves will yield extraneous intersections.

So the theorem will hold true, for isosceles triangles, if we could modify the hypothesis variety, restricting to one specific component (the one that corresponds to the interior incenter). Now, finding the equations for one such component implies, in general, performing decomposition methods, and it leads to much more elaborated, less performing, algorithms, far from the present simplicity of our approach. Nevertheless, at least we can say the method succeeded giving us the intuition to discover what was 'really' true, even if it could not be proved by the same technique.

3. Conclusions

The novelty and interest of our present work could be that rather than focusing on proving or disproving theorems, our goal is on finding complementary hypothesis for a(ny) given conjectural statement to hold true. Although incomplete (if factorization of varieties is to be avoided) when dealing with statements that are not generally true and not generally false, it is in practice quite performing (see [13]).

As the automatic discovery method provides means to explore open geometric situations, it can have many applications. Among other, the proposed method could be regarded as the core of a future program—dedicated to elementary teaching—that allows, when linked simultaneously with a tool for displaying geometric constructions and a symbolic computation package, the interactive exploration of geometric properties.

As is well known, Cabri-Géomètre provides an interactive learning environment for elementary geometry; in particular it has been considered as an instrument for theorem proving in this context [11]. We think that the interaction of such tool with our method could provide an intelligent, interactive environment for learning Euclidean geometry. The idea is to build a sort of *Geometry Guide* program that not only allows one to experiment properties and to display figures, but that also has the ability of 'knowing' in advance what is the correct direction that the user has to follow if some geometrical construction is set, in case he/she wants to find some stated property (or, using automatic derivation techniques, what could be some interesting properties hidden in a given construction). Obviously, for this particular application, understanding how the program itself proceeds should not be considered as a requirement for the child-user; the core of the discovery method should be a kind of black box for that particular user. Maybe this is too much of a future dream, as the present state of algebraic computation technology is too far from providing a sound basis for attacking too complicated geometrical situations (in particular, since we are lacking of a sufficiently developed 'real' geometry approach). But it is, anyhow, a promising research idea, in our opinion.

A different didactical application to Algebraic Geometry teaching at higher education level is discussed in [14].

References

1. Chou, S. C.: *Mechanical Geometry Theorem Proving*, D. Reidel, 1987.
2. Chou, S. S. and Gao, X. S.: Methods for mechanical geometry formula deriving, in *Proceedings ISSAC-90*, ACM, New York, 1990, pp. 265–270.
3. Conti, P. and Traverso, C.: A case of automatic theorem proving in Euclidean geometry: The Maclane 8_3 theorem, in *Proceedings AAECC*, Paris, 1995.

4. Cox, D., Little, J. and O'Shea, D.: *Ideals, Varieties and Algorithms*, Undergraduate Texts in Mathematics, Springer-Verlag, 1992.
5. Guergueb, A.: *Examples de demostration automatique en gèométrie réelle*, Ph. Dissertation, Univ. de Rennes I, 1994.
6. Kapur, D.: Geometry Theorem Proving Using Hilbert's Nullstellensatz, in B. W. Char (ed.), *Proc. of the 1986 Symposium on Symbolic and Algebraic Computation*, Waterloo, Ont., 1986.
7. Kapur, D.: A refutational approach to theorem proving in geometry, *Artif. Intell.* **37**(1–3) (1988), 61–93.
8. Kapur, D. and Mundy, J. L.: Wu's method and its application to perspective viewing, in D. Kapur and J. L. Mundy (eds.), *Geometric Reasoning*, The MIT press, Cambridge, Mass., 1989.
9. Koepf, W.: Gröbner bases and triangles. *Internat. J. Comput. Algebra in Math. Education* 4(4) (1998), 371–386.
10. Kutzler, B. and Stifter, S.: Automated geometry theorem proving using Buchbergers algorithm, in B. W. Char (ed.), *Proc. of the 1986 Symposium on Symbolic and Algebraic Computation*, Waterloo, Ont., 1986.
11. Laborde, J. M. and Strässer, R.: Cabri géomètre, a microworld of geometry for guided discovery learning, *Zentralblatt für Didaktik der Mathematik* **90**(5), 171–190.
12. Mainguené, J.: *Méthode de Wu, courbes réelles et demostration automatique en Gèométrie*, Ph. Dissertation, Univ. de Rennes I, 1994.
13. Recio, T.: *Cálculo Simbólico y Geométrico*, Editorial Síntesis, Madrid, 1998.
14. Recio, T., Sterk, H. and Vélez, M. P.: Automatic geometry theorem proving, in A. M. Cohen, H. Cuipers, and H. Sterk (eds.), *Some Tapas of Computer Algebra*, Algorithms and Computations in Math. 4, Springer-Verlag, 1998.
15. Wang, D.: Gröbner bases applied to geometric theorem proving and discovering, in B. Buchberger and F. Winkler (eds.), *Gröbner Bases and Applications*, London Math. Soc. Lecture Note Series, Cambridge University Press, Cambridge, 1998.
16. Wu, W. T.: On the decision problem and the mechanization of theorem-proving in elementary geometry, *Scientia Sinica* **21**, 159–172 (also in W. W. Bledsoe and D. W. Loveland (eds.), *Automated Theorem Proving: After 25 Years*, American Mathematical Society, Providence, 1984, pp. 213–234).
17. Wu, W. T.: *Mechanical Theorem Proving in Geometries*, Texts and Monographs in Symbolic Computation, Springer-Verlag, Wien, 1994.