# Automatic discovery of geometry theorems using minimal canonical comprehensive Gröbner systems

Antonio Montes<sup>\*</sup>, Tomas Recio<sup>†</sup>

#### Abstract

The main idea in this paper is merging two techniques that have been recently developed. On the one hand, we consider MCCGS, standing for Minimal Canonical Comprehensive Groebner Systems, a recently introduced computational tool yielding "good" bases for ideals of polynomials over a field depending on several parameters, that specialize "well", for instance, regarding the number of solutions for the given ideal, for different values of the parameters. The second ingredient concerns automatic theorem discovery in elementary geometry. Automatic discovery aims to obtain complementary hypotheses for a (generally false) geometric statement to become true. The paper shows how to use MCCGS for automatic discovering of theorems and gives relevant examples.

*Key words:* automatic discovering, comprehensive Gröbner system, automatic theorem proving, canonical Gröbner system, elementary geometry. *MSC:* 14Q99, 51N20, 68T15.

### 1 Introduction

The main idea in this paper is the merging of two techniques that have been recently developed. On the one hand, we will consider a recent proposal (named *MCCGS*, standing for *minimal canonical comprehensive Gröbner systems*) [MaMo06], that is –roughly speaking– a computational tool yielding "good" bases for ideals of polynomials over a field depending on several parameters, where "good" means that the obtained bases should specialize (and specialize "well", for instance, regarding the number of solutions for the given ideal) for different values of the parameters.

Briefly, in order to understand what kind of problem MCCGS addresses, let us consider the ideal  $(ax, x + y)\mathbb{Q}[a][x, y]$ , where a is taken as a parameter. Then it is clear that there will be different bases for the specialized ideal  $(a_0x, x+y)\mathbb{Q}[x, y]$ , one for  $a_0 = 0$  and another one for rational values such that  $a_0 \neq 0$  (in the former case (x + y) is a Gröbner-basis (in short, a G-basis) for the specialized ideal; in the latter case, a G-basis will be (x, y)). On the other hand, let us consider  $(ax - b)\mathbb{Q}[a, b][x]$ , where a, b are taken as free parameters and x is the only variable. Then, no matter which rational values  $a_0, b_0$  are assigned to a, b,

<sup>\*</sup>Work partially supported by the Spanish Ministerio de Ciencia y Tecnología under project MTM 2006-01267, and by the Generalitat de Catalunya under project 2005 SGR 00692.

<sup>&</sup>lt;sup>†</sup>Work partially supported by the Spanish Ministerio de Educación y Ciencia under project GARACS, MTM2005-08690-C02-02.



Figure 1: Orthic triangle

it happens that  $\{a_0x - b_0\}$  remains a Gröbner basis for  $(a_0x - b_0)\mathbb{Q}[x]$ . Still, there is a need for a case-distinction if we focus on the cardinal of the solutions for the specialized ideal. Namely, for  $a_0 \neq 0$  there is a unique solution  $x = -b_0/a_0$ ; for  $a_0 = 0$  and  $b_0 \neq 0$  there is no solution at all; and for  $a_0 = b_0 = 0$  a solution can be any value of x (no restriction, one degree of freedom).

The goal of *MCCGS* is to describe, in a compact and canonical form, the discussion, depending on the different values of the parameters specializing a given parametric system, on the different kind of systems and their solutions.

The second ingredient of our contribution is about automatic theorem discovery in elementary geometry. Automatic discovery aims to obtain complementary hypotheses for a (generally false) geometric statement to become true. For instance, we can consider an arbitrary triangle and the feet on each sides of the three altitudes. These three feet give us another triangle, and now we want to conclude that such triangle is equilateral. This is generally false, but, under what extra hypotheses on the given triangle will it become true?

Finding, in an automatic way, the necessary and sufficient conditions for this statement to become a theorem, is the task of automatic discovery.

Our goal in this paper is to show how performing a *MCCGS* procedure on a certain ideal built up from the given hypotheses and thesis, depending on the free coordinates of some elements of the geometric setting, can improve the automatic discovery of geometry theorems.

This idea is related to the work of [CLLW], inspired by [K95] and by [Weis92]. In that paper, a parametric radical membership test is presented for a mathematical construct the authors introduce, called partitioned parametric Gröbner basis. We notice the authors of [CLLW] mention the paper of Montes [Mo02] as being a predecessor of their work on the discussion of Gröbner basis with parameters. It turns that the parametric radical membership test gives, in a straightforward manner, when applied to the ideal  $(h_1 \dots h_r, gy - 1)$  of hypotheses  $h_1 \dots h_r \in K[u, x]$  (for some field K) plus the negation of the thesis  $g \in K[u, x]$ , the collection of all non-degeneracy conditions required for proving such theorem, with the conditions being expressed in terms of the free parameters u of the geometric situation.

Roughly speaking, the partitioned basis of an ideal  $I \subseteq K[u, x]$  is a finite collection

of couples  $(C_i, F_i)$ , where the  $C_i$ 's are constructible sets on the parameter space, and the  $F_i$ 's are G-bases in K[u, x]. Moreover, it is required that the  $C_i$ 's conform a partition of the parameter space and, also, that for every element  $u_0$  in each  $C_i$ , the G-basis of  $(h_1(u_0, x) \dots h_r, (u_0, x), g(u_0, x) y - 1)$  is precisely  $F_i(u_0, x)$ . It is well known (e.g. [K86] or [Ch88]) that, in this context, a theorem  $\{h_1 = 0 \dots h_r = 0\} \implies \{g = 0\}$  is to be considered true if  $1 \in (h_1 \dots h_r, gy - 1)$ ; thus the non-degeneracy conditions are precisely those expressed by the  $C_i$ 's such that  $F_i = \{1\}$ , since this is the only case  $F_i$  can specialize to  $\{1\}$ . Yet we must remark that, simply testing for  $1 \in (h_1 \dots h_r, gy - 1)$ , can yield theorems that hold just because the hypotheses are not compatible (i.e. such that already  $1 \in (h_1 \dots h_r)$ ). This cannot happen with our approach to automatic discovery (see next Section ): if a new statement is discovered, then the obtained hypotheses will be necessarily compatible.

Our contribution differs from [CLLW] in two senses: first, we focus on automatic discovery, and not in automatic proving. Second, the use of MCCGS provides not only the specialization property (which is the key for the application of partitioned parametric bases in [CLLW]) but also a case distinction, that allows a richer understanding of the underlying geometry for the considered situation. In fact, it seems that the partitioned parametric G-Basis (PPGB) algorithm from [CLLW] is close to the algorithm DISPGB considered in [Mo02], both sharing that their output requires collecting by hand multiple cases where  $F_i = \{1\}$  (and then having to manually express in some simplified way the union of the corresponding conditions on the parameters). Actually, the motivation for MCCGS was, precisely, improving DISPGB.

Next Section includes a short introduction to the basics on automatic proving, exemplified in Section 3 via the more traditional way. Section 4 provides some bibliographic references for the problem of the G-basis specialization and summarizes the main features of the MCCGS algorithm, including an example of its output. Section 5 describes the application of MCCGS to automatic discovery, together with a collection of curious examples, including the solution of a pastime from Le Monde and the simpler solution (via this new method) of the previous example from Section 3.

# 2 A digest on automatic discovery

As mentioned above, automatic discovery aims to obtain complementary hypotheses for a (generally false) geometric statement to become true, such as stating that the three feet of the altitudes for a given triangle form themselves an equilateral triangle.

Even if less popular than automatic proving, automatic discovery of elementary geometry theorems is not new. It can be traced back to the work of Chou (see [Ch84], [Ch87] and [ChG90]), generally as a task for "automatic derivation of formulas", a kind of automatic discovery in which the conjectured thesis is a trivial statement such as 0 = 0. Then, the search for complementary hypotheses for that thesis to hold, consists in deriving results that always occur under the given hypotheses, but restricted to searching those results formulated in terms of some specific set of variables (such as expressing the area of a triangle in terms of the lengths of its sides).

Further specific contributions to automatic discovery appear in [K89], [Wa98], [R98]

(a book written in Spanish for secondary education teachers, with circa one hundred pages devoted to this topic and with many worked out examples), [RV99], [Ko] or [CW]. Examples achieved through a specific software for discovery, named *GDI* (the initials of *Geometría Dinámica Inteligente*), of Botana-Valcarce, appear in [BR05] or [RB], such as the automatic derivation of the thesis for the celebrated Maclane 8<sub>3</sub>-Theorem, or the automatic approach to some items on a test posed by Richard [Ri], on proof strategies in mathematics courses, for students 14-16 years old.

The simple idea behind the different approaches is, essentially, that of adding the conjectural thesis to the collection of hypotheses, and then deriving, from this new ideal of thesis plus hypotheses, some new constraints in terms of the free parameters ruling the geometric situation. For a toy example, consider that x - a = 0 is the only hypothesis, where a is a parameter, and that x = 0 is the (generally false) thesis. Then we consider (x - a, x) as the new ideal, which contains the constraint a = 0, and this is indeed the extra condition we have to add to x - a = 0, in order to verify the thesis x = 0. A detailed description of the procedure we will follow for automatic discovery appears in [RV99], and it has been recently revised in [BDR] and [DR], showing that, in some precise sense, the procedure is intrinsically unique.

Let us recall here that the approach to discovery in [RV99] proceeds, roughly speaking, first identifying a set of independent variables (those ruling the construction of the hypothesis variety, defined by the zeroes over an algebraically closed field of the hypotheses ideal H). The corresponding components of this variety, where these variables are independent (and maximally independent, as well) are called *privileged*. Let us consider the *Saturation* (e.f. [KR00]) of the ideal of hypotheses by the ideal T of thesis,  $H : T^{\infty}$ , according to the following

**Definition 1.** Take I, J ideals of K[X]. Recall that  $I : J = \{x, xJ \subset I\}$ . Then, the saturation of I by J is defined as  $I : J^{\infty} = \bigcup_n (I : J^n)$ .

Notice the saturation of I by J gives the intersection of all primary components q associated to prime ideals of a minimal decomposition of I such that there is an f in J with f not in such primes, i.e. the saturation of I by J is the intersection of the primary components associated to the primes such that J is not contained in them.

With this notation it can be shown (see [RV99], [DR]) that the elimination ideal (over the independent variables), of  $H: T^{\infty}$  is not zero if and only if the theorem is true over all the privileged components (and then the theorem is called "generally true" [Ch88]).

When the given theorem is not generally true, it turns that the elimination ideal of the ideal generated by the hypotheses plus the thesis is not zero if and only if the thesis does not hold over any privileged component (the so called "generally false" case, the one suitable for discovery).

In this latter situation, [RV99] considers adding, as new hypotheses, the equations provided by the elimination of the old hypothesis plus the thesis, and proceeds further on, identifying a subset of the privileged variables that remain maximally independent over the new hypothesis variety.

This new set, the union of the hypotheses and the given thesis , yields a non-generally false theorem, and, in many interesting examples, it is generally true (but not always: the



Figure 2: Problem of Example 2

method is incomplete without introducing factorization, as shown in [RV99]; see also the last comment on [CLLW] ).

### 3 An example

Next, we will develop the above introduced notions considering a statement from [Ch88] (Example 91 in his book), suitably adapted to the discovery framework. The example here is taken from [DR].

**Example 2.** Let us consider as given data a circle and two diametral opposed points on it (say, take a circle centered at (1,0) with radius 1, and let C = (0,0), D = (2,0) the two ends of a diameter), plus an arbitrary point  $A = (u_1, u_2)$ . See Figure 2. Then trace a tangent from A to the circle and let  $E = (x_1, x_2)$  be the tangency point. Let  $F = (x_3, x_4)$  be the intersection of DE and CA. Then we claim that AE = AF. Moreover, in order to be able to define the lines DE, CA, we require, as hypotheses, that  $D \neq E$  (ie.  $u_1 \neq 2$ ) and that  $C \neq A$  (ie.  $u_1 \neq 0$  or  $u_2 \neq 0$ ).

Now, using CoCoA [CNR99] and its package TP (for Theorem Proving), we translate the given situation as follows

Alias TP := \$contrib/thmproving;

```
Use R::=Q[x[1..4],u[1..2]];
```

```
A:=[u[1],u[2]];
E:=[x[1],x[2]];
D:=[2,0];
```

```
F:=[x[3],x[4]];
C:=[0,0];
Ip1:=TP.Perpendicular([E,A],[E,[1,0]]);
Ip2:=TP.LenSquare([E,[1,0]])-1;
Ip3:=TP.Collinear([0,0],A,F);
Ip4:=TP.Collinear(D,E,F);
```

T:=Ideal(TP.LenSquare([A,E])-TP.LenSquare([A,F]));

where T is the thesis and H describes the hypothesis ideal. Notice that Ip1 expresses that the segments [E, A], [E, (1, 0)] are perpendicular; Ip2 states that the square of the length of [E, (1, 0)] is 1 (so Ip1, Ip2 imply E is the tangency point from A); and the next two hypotheses express that the corresponding three points are collinear. The hypothesis ideal H is here constructed by using the *saturation* command, since it is a standard way of stating that the hypothesis variety is the (Zariski) closure of the set defined by all the conditions Ip[i] = 0, i = 1...4 minus the union  $\{u[1] = 2\} \cup \{u[1] = 0, u[2] = 0\}$ , as declared in the formulation of this example (but we refer to [DR] for a discussion on the two possible ways of introducing inequalities as hypotheses). Finally, the thesis expresses that the two segments [AE], [AF] have equal non oriented length.

First we check that the statement  $H \implies T$  is not algebraically true in any conceivable way. For instance, it turns that

Saturation(H, Saturation(H,T)); Ideal(1)

and this computation shows that all possible non-degeneracy conditions (those polynomials  $p(\mathbf{u}, \mathbf{x})$  that could be added to the hypotheses as conditions of the kind  $p(\mathbf{u}, \mathbf{x}) \neq \mathbf{0}$ ) lie in the hypothesis ideal, yielding, therefore to an empty set of conditions of the kind  $p \neq 0 \land p = 0$ . This implies, in particular, that the same negative result would be obtained if we restrict the computations to some subset of variables, since the thesis does not vanish on any irreducible component of the hypothesis variety.

Thus we must switch on to the discovery protocol, checking before hand that u[1], u[2] actually is a (maximal) set of independent variables –the parameters– for our construction:

Dim(R/H);
2
-----Elim([x[1],x[2],x[3],x[4]],H);
Ideal(0)

Then we add the thesis to the hypothesis ideal and we eliminate all variables except u[1], u[2]

yielding as complementary hypotheses the conditions  $u[1]^2 + u[2]^2 - 2u[1] = 0 \lor u[1] = 0$ that can be interpreted by saying that either point A lies on the given circle or (when u[1] = 0) triangle  $\Delta(A, C, D)$  is rectangle at C. In the next step of the discovery procedure we consider as new hypothesis ideal the set H + H', which is of dimension 1 and where both u[2] or u[1] can be taken as independent variables ruling the new construction.

Dim(R/(H+H'));
1
----Elim([x[1],x[2],x[3],x[4],u[1]],H+H');
Ideal(0)
----Elim([x[1],x[2],x[3],x[4],u[2]],H+H');
Ideal(0)

Choosing, for example, u[2] as relevant variable, we check –applying the usual automatic proving scheme– that the new statement  $H \wedge H' \implies T$  is correct under the non-degeneracy condition  $u[2] \neq 0$ :

```
H'':=Elim([x[1],x[2],x[3],x[4],u[1]], Saturation(H+H',T));
H'';
Ideal(u[2]^3)
```

Thus we have arrived to the following statement: Given a circle of radius 1 and centered at (1,0), and a point A not in the X-axis and lying either on the Y axis or in the circle, it holds that the segments AE, AF (where E is a tangency point from A to the circle and F is the intersection of the lines passing by (2,0), E and A, (0,0)) are of equal length.

# 4 Overwiew on the MCCGS algorithm

As mentioned in the introduction, specializing the basis of an ideal with parameters does not yield, in general, a basis of the specialized ideal.

This phenomenon –in the context of Gröbner basis– has been known for over fifteen years now, yielding to a rich variety of attempts towards a solution (we refer the interested reader to the bibliographic references in [MaMo06] or in [Wib06]). Finding a specializable basis (ie. providing a single basis that collects all possible bases, together with the corresponding relations among the parameters) is -more or less- the task of the different comprehensive G-Basis proposals. Although the first global solution was that of Weispfenning, as early as 1992 (see [Weis92]), the topic is quite active nowadays, as exemplified in the above quoted recent papers. The *MCCGS* procedure, that is, computing the *minimal canonical comprehensive Gröbner system* of a given parametric ideal, is one of the approaches we are interested in. Let us describe briefly the goals and output of the *MCCGS* algorithm.

Given a parametric polynomial system of equations, our interest focuses on discussing the type of solutions depending on the values of the parameters. Let  $\mathbf{x} = (x_1, \ldots, x_n)$  be the set of variables,  $\mathbf{u} = (u_1, \ldots, u_m)$  the set of parameters and  $I \subset \mathbb{Q}[\mathbf{u}][\mathbf{x}]$  the parametric ideal we want to discuss. We want to study how the complex solutions of the equation system defined by I vary when we specialize the values of the parameters  $\mathbf{u}$  to concrete values  $\alpha \in \mathbb{C}$ . Denote by  $A = \mathbb{Q}[\mathbf{u}]$ , and by  $\sigma_{\mathbf{u}_0} : A[\mathbf{x}] \to \mathbb{C}[\mathbf{x}]$  the homomorphism corresponding to the specialization (substitution of  $\mathbf{u}$  by some  $\mathbf{u}_0 \in \mathbb{C}$ ).

A Gröbner System of the ideal  $I \subset A[\mathbf{x}]$  wrt (with respect to) the termorder  $\succ_{\mathbf{x}}$  is a set

$$GS(I, \succ_{\mathbf{x}}) = \{ (S_i, B_i) : 1 \le i \le s, S_i \subset \mathbb{C}^m, B_i \subset A[\mathbf{x}], \bigcup_i S_i = \mathbb{C}^m, \\ \forall \mathbf{u}_0 \in S_i, \sigma_{\mathbf{u}_0}(B_i) \text{ is a Gröbner basis of } \sigma_{\mathbf{u}_0}(I) \text{ wrt } \succ_{\mathbf{x}} \}.$$

The algorithm MCCGS (Minimal Canonical Comprehensive Gröbner System) [Mo06, MaMo06] of the ideal  $I \subset A[\mathbf{x}]$  wrt the monomial order  $\succ_{\mathbf{x}}$  for the variables, builds up the unique Gröbner System having the following properties:

- 1.  $S = \{S_1, \ldots, S_s\}$  is a partition of the parameter space  $\mathbb{C}^m$ .
- 2. The bases  $B_i$  are normalized to have content 1 wrt **x** over  $\mathbb{Q}[\mathbf{u}]$  (in order to work with polynomials instead of with rational functions), and the leading coefficients are different from zero on every point of  $S_i$ . Moreover, the  $B_i$  specialize to the reduced Gröbner basis of  $\sigma_{\mathbf{u}_0}(I)$ , keeping the same lpp's (leading power products set) for each  $\mathbf{u}_0 \in S_i$ . Thus a concrete set of lpp's can be associated to a given  $S_i$ . Moreover, although a same set of lpp's can be attached to different  $S_i$ 's, if two segments  $S_i, S_j$ share the same lpp's, then there is not a common basis B specializing to both  $B_i, B_j$ .
- 3. The partition  $\mathcal{S}$  is canonical (unique for a given I and monomial order).
- 4. The partition is minimal, in the sense it does not exists another partition having property 2 with less sets  $S_i$ .
- 5. The sets  $S_i$  (often called *segments*) are constructible and are described in a canonical form.

As it is known, the lpp's of the reduced Gröbner basis of an ideal determine the cardinal or dimension of the solution set over an algebraically closed field. This makes the *MCCGS* algorithm very useful for applications as it identifies canonically the different kind of solutions for every value of the parameters. This is particularly suitable for automatic theorem proving and automatic theorem predicting, as we will show in the following sections.

Let us give an example of the output of MCCGS.

**Example 3.** Consider the system described by the following parametric ideal (here the parameters are a, b, c, d):

$$I = (x^{2} + by^{2} + 2cxy + 2dx, 2x + 2cy + 2d, 2by + 2cx),$$

arising in the context of finding all possible singular conics and their singularities. Calling to the Maple implementation of MCCGS yields a graphical and an algebraic output. The graphical output is shown in Figure 3. It contains the basic information that is to be read as follows. At the root there is the given ideal (in red). The second level (also in red) contains the lpp's of the bases of the three different possible cases. These are [1], corresponding to the no solution (no singular points) case; [x, y], corresponding to the one solution (one singular point) case; and [x], corresponding to the case of one dimensional solution (ie. when the conic is a double line). Below each case there is a subtree (in blue) describing the corresponding  $S_i$ , with the following conventions:

- at the nodes there are prime ideals of  $\mathbb{Q}[\mathbf{u}]$ ,
- a descending edge means the set theoretic "difference" of the set defined by the node above minus the set defined at the node below,
- nodes at the same level, hanging from a common node, are to be interpreted as yielding the set theoretic "union" of the corresponding sets,
- every branch contains a strictly ascending chain of prime ideals.

So, in the example above, the three cases, their lpp's and the corresponding  $S_i$ 's are to be read as shown in the following table:

lpp	Basis $B_i$	Description of $S_i$
[1]	[1]	$\mathbb{C}^3 \setminus \left( (\mathbb{V}(b) \setminus (\mathbb{V}(c,b) \setminus \mathbb{V}(d,c,b)) \right) \cup \mathbb{V}(d) \right)$
[y,x]	[2cy+d,x]	$(\mathbb{V}(b) \setminus \mathbb{V}(c,b)) \cup (\mathbb{V}(d) \setminus \mathbb{V}(d,b-c^2))$
[x]	[x+cy]	$\mathbb{V}(d,b-c^2)$

We remark that the  $B_i$ 's do not appear in the Figure 3, since –in order to simplify the display– the complete bases are only given by the algebraic output of *MCCGS* and are not shown by the graphic output.

## 5 Using *MCCGS* for automatic theorem discovering

Once we have briefly described the context for *MCCGS* and for automatic discovery, we are prepared to describe the basic idea in this paper. We can say that our goal is to show how performing a *MCCGS* procedure can improve the automatic discovery of geometry theorems.

Example 3 can be seen as a very simple example of theorem discovering. We could formulate the statement a conic has one singular point and try to find the conditions for



Figure 3: MCCGS for the singular points of a conic

the statement to be true. Without loss of generality we express the equation of the conic and its partial derivatives as

$$I = (x^{2} + by^{2} + 2cxy + 2dx, 2x + 2cy + 2d, 2by + 2cx),$$

and search for the values of the parameters where this system has a single solution. As shown above, we have found that the statement is true if and only if  $\{b = 0, c \neq 0\}$  or if  $\{d = 0, b - c^2 \neq 0\}$ , since in the first case there is no solution  $(B_1 = (1))$ , while the third case yields a 1-dimensional set of solutions.

As stated in sections 2 and 3, automatic discovery can be approached considering as new hypotheses the generators of the elimination, over the free parameters of the problem, of the ideal of hypotheses and thesis I = (H, T). This is, precisely, the (Zariski closure of the) projection, over the parameter space, of the zero set of this ideal I. It is clear, then, that, for automatic discovery through MCCGS one must perform such decomposition over I, selecting those segments  $S_i$  such that the system I has at least one solution (in the complex field or in whatever algebraically closed field we are working with) over  $S_i$ . In other words, discarding the  $S_i$ 's with  $B_i$  equal to 1 and keeping the remaining  $S_i$ 's. The description of these  $S_i$ 's gives, precisely, the new conditions for the thesis to hold over the hypotheses variety.

Let us see how this works in a collection of examples, where we have just detailed the discovery step in the procedure outlined above. That is, we have not included here the verification in each case that the newly found hypotheses actually lead to a true statement (the proving step, which should be performed in the standard way; in particular, it could be done using MCCGS to test if 1 belongs to the saturation of the ideal of new hypotheses by the thesis).

**Example 4.** (See also [DR]). Let us now review Example 2 using *MCCGS*. As seen there, the hypothesis are the union of  $H := H_1 \cup S$ , where  $H_1$  expresses the equality type constraints:

$$H_1 = [(x_1 - 1)(u_1 - x_1) + x_2(u_2 - x_2), (x_1 - 1)^2 + x_2^2 - 1, u_1x_4 - u_2x_3, x_3x_2 - x_4x_1 - 2x_2 + 2x_4]$$

to which we have to add the saturation ideal expressing the inequality constraints:

$$S = \begin{bmatrix} u_1x_4 - u_2x_3, x_1u_1 - u_1 - x_1 + x_2u_2, x_4x_2 - 2x_2u_2 - x_3u_1 + 2u_1, \\ x_4x_1 - 2x_1u_2 + u_2x_3, x_3x_2 - 2x_1u_2 + u_2x_3 - 2x_2 + 2x_4, \\ x_1x_3 + x_3u_1 + 2x_2u_2 - 2x_1 - 2u_1, x_1^2 - 2x_1 + x_2^2, \\ x_3u_1^2 + 2x_2u_2u_1 - 2u_2^2x_1 + u_2^2x_3 - 2u_1^2 - 2x_2u_2 + 2u_2x_4, \\ x_3^2u_1 + x_4u_2x_3 + 2x_4^2 - 4x_3u_1 - 4u_2x_4 + 4u_1, \\ u_1x_2^2 - x_1x_2u_2 - x_2^2 + x_2u_2 + x_1 - u_1, \\ u_2x_3^3 + u_2x_4^2x_3 + 2x_4^3 - 4u_2x_3^2 - 4u_2x_4^2 + 4u_2x_3]. \end{bmatrix}$$

The thesis is

$$T = (u_1 - x_1)^2 + (u_2 - x_2)^2 - (u_1 - x_3)^2 - (u_2 - x_4)^2.$$

Calling now

$$mccgs(H_1 \cup S \cup T, lex(x_1, x_2, x_3, x_4), lex(u_1, u_2))$$

one obtains the following segments:

Segment	lpp	Description of $S_i$
1	[1]	$\mathbb{C}^2 \setminus (\mathbb{V}(u_1^2 + u_2^2 - 2u_1) \cup \mathbb{V}(u_1))$
2	$[x_4^2, x_3, x_2, x_1]$	$\mathbb{V}(u_1^2 + u_2^2 - 2u_1) \setminus (\mathbb{V}(u_1 - 2, u_2) \cup \mathbb{V}(u_1, u_2))$
3	$[x_4^2, x_3, x_2, x_1]$	$\mathbb{V}(u_1) \setminus (\mathbb{V}(u_1, u_2^2 + 1) \cup \mathbb{V}(u_1, u_2))$
4	$[x_4, x_3, x_2, x_1]$	$\mathbb{V}(u_1, u_2^2 + 1)$
5	$[x_4^2, x_3, x_2^2, x_1]$	$\mathbb{V}(u_1-2,u_2)$
6	$[x_4^2, x_3^2, x_2, x_1]$	$\mathbb{V}(u_2, u_1)$

Segment  $S_1$  states that point  $A(u_1, u_2)$  must lie either in the Y-axis or on the circle, as a necessary condition in the parameter space  $\mathbf{u} = (u_1, u_2)$  for the existence of solutions, in the hypothesis plus thesis variety, lying over  $\mathbf{u}$ . This essentially agrees with the result obtained in section 2.

A detailed analysis of the remaining segments show a variety of formulas for determining the (sometimes not unique) values of points  $E(x_1, x_2)$  and  $F(x_3, x_4)$  -verifying the theoremover the corresponding parameter values.

For completeness we give the different bases associated, in the different segments, to the above ideal of thesis plus hypotheses

$$\begin{array}{rcl} B_1 &=& [1] \\ B_2 &=& [u_2^2 + x_4^2 - 2u_2x_4, -u_1x_4 + u_2x_3, u_2^3 - 2u_2u_1 + x_2u_2^2 + (-2u_2^2 + 2u_1)x_4, \\ && u_2u_1 + x_1u_2 - 2u_1x_4] \\ B_3 &=& [-2u_2x_4 + x_4^2, x_3, (u_2^2 + 1)x_2 - x_4, (u_2^2 + 1)x_1 - u_2x_4] \\ B_4 &=& [x_4, x_3, x_2, x_1] \\ B_5 &=& [x_4, -4 + 2x_3, x_2^2, -2 + x_1] \\ B_6 &=& [x_4^2, x_3^2, -x_3x_4 + 2x_2 - 2x_4, 2x_1] \end{array}$$

**Example 5.** Next we consider the problem described in Figure 4<sup>1</sup>. Take a circle C with center at O(0,0) and radius 1 and let us denote points A = (-1,0) and B = (0,1). Let D

<sup>&</sup>lt;sup>1</sup>We thankfully acknowledge here that this problem was suggested by a colleague, Manel Udina



Figure 4: Example 5.

be an arbitrary point with coordinates D = (1+a, b) and let C = (1+a, 0) be another point in the X-axis, lying under point D. Then trace the line BC. Assume this line intersects the circle C at point P(x, y).

Consider now the, in general false, statement "the points A, P, D are aligned". We want to discover the conditions on the parameters a, b for the statement to be true. The set of hypothesis plus thesis equations are very simple:

$$HT = [x^{2} + y^{2} - 1, -x + 1 - y + a - ay, -2y + b + xb - ay]$$

Take x, y as variables and a, b as parameters and call mccgs(HT, lex(x, y), lex(a, b)). The graphical output of the algorithm can be seen in Figure 5, and the algebraic description appears in the following table.

lpp	Basis $B_i$	Description of $S_i$
[1]	[1]	$(\mathbb{C}^2 \setminus (\mathbb{V}(a-b) \setminus \mathbb{V}(a-b,(b+1)^2+1)))$
		$\cup (\mathbb{C}^2 \setminus \mathbb{V}(2+a))$
		$\cup (\mathbb{C}^2 \setminus \mathbb{V}(a-b+2))$
[y,x]	$[x^2 + y^2 - 1,$	$(\mathbb{V}(a-b) \setminus \mathbb{V}(a-b,(b+1)^2+1))$
	x + (a+1)(y-1),	$\cup \left(\mathbb{V}(2+a) \setminus \mathbb{V}(b,2+a)\right)$
	b(x+1) - (a+2)y]	$\cup (\mathbb{V}(a-b+2) \setminus \mathbb{V}(b,2+a))$
$[y^2, x]$	[y(y-1), 1+x-y]	$\mathbb{V}(b, 2+a)$

As we see, the generic case has basis [1] showing that the statement is false in general. The interesting case corresponds, as it is usually expected, to the case with lpp = [x, y], providing a unique solution for P. The description of the parameter set associated to this



Figure 5: Canonical tree for Example 5

basis gives the union of three different locally closed sets, namely  $\mathbb{V}(a-b)\setminus\mathbb{V}(a-b,(b+1)^2+1)$ ,  $\mathbb{V}(2+a)\setminus\mathbb{V}(b,2+a)$  and  $\mathbb{V}(a-b+2)\setminus\mathbb{V}(b,2+a)$ , expressing complementary hypotheses for the statement to hold.

The first set is (perhaps) the expected one, corresponding to the case a = b (except for the degenerate complex point (b, b) with  $(b + 1)^2 + 1) = 0$ , without interest from the real point of view). Thus we can say that the statement holds if point C is equidistant from point D and point E.

The second set yields a = -2 and corresponds to the situation where point D is on the tangent to the circle trough the point (-1,0) (except for the degenerate case b = 0). In this case P = A and, obviously, A, P, D are aligned (even in the degenerate case, as stated in the third segment, corresponding to the  $lpp[y^2, x]$ ).

Finally, the third set gives the condition b = a + 2 and it is also interesting, since it corresponds to the case where the intersecting point of the line BC with the circle is taken to be B instead of P, and then point D' should be in the vertical of C and at distance D'C equal to distance EC plus two.

#### **Example 6.** [Isosceles orthic triangle]

In [DR] the conditions for the orthic triangle of a given triangle (that is, the triangle built up by the feet of the altitudes of the given triangle over each side) to the equilateral have been discovered. Next example aims to discover conditions for a given triangle in order to have an isosceles orthic triangle.

Consider the triangle of Figure 1 with vertices A(-1,0), B(1,0) and C(a,b), corresponding to a generic triangle having one side of length 2. Denote by  $P_1(a,0)$ ,  $P_2(x_2,y_2)$ ,  $P_3(x_3,y_3)$  the feet of the altitudes of the given triangle, i.e. the vertices of the orthic triangle. The equations defining these vertices are:

$$H = (a-1)y_2 - b(x_2 - 1) = 0,$$
  

$$(a-1)(x_2 + 1) + by_2 = 0,$$
  

$$(a+1)y_3 - b(x_3 + 1) = 0,$$
  

$$(a+1)(x_3 - 1) + by_3 = 0,$$



Figure 6: Canonical tree branch for  $lpp = [y_3, y_2, x_3, x_2]$  in Example 6.

Now let us add the condition  $\overline{P_1P_3} = \overline{P_1P_2}$ .

$$T = (x_3 - a)^2 + y_3^2 - (x_2 - a)^2 - y_2^2 = 0.$$

Take  $x_2, x_3, y_2, y_3$  as variables and a, b as free parameters and call

 $\operatorname{mccgs}(H \cup T, \operatorname{lex}(x_2, x_3, y_2, y_3), \operatorname{lex}(a, b)).$ 

The output has now four segments. The generic case, with lpp = [1], meaning that the orthic triangle is, in general, not isosceles; one interesting case with  $lpp = [y_3, y_2, x_3, x_2]$ ; and two more cases we can call degenerate, with lpp's  $[y_2, x_3^2, x_2]$  and  $[y_2, x_3, x_2^2]$ , respectively. For the interesting case we show the graphic output in Figure 6. Its basis is

$$B_2 = [(a^2 + b^2 + 2a + 1)y_3 - 2ab - 2b, (a^2 + b^2 - 2a + 1)y_2 + 2ab - 2b, (a^2 + b^2 + 2a + 1)x_3 - a^2 + b^2 - 2a - 1, (a^2 + b^2 - 2a + 1)x_2 + a^2 - b^2 - 2a + 1].$$

Next table shows the description of the lpp and the  $S_i$ 's for the the four cases:

lpp	Description of $S_i$
[1]	$\mathbb{C}^2 \setminus \left( (\mathbb{V}(a) \setminus \mathbb{V}(b^2 + 1, a)) \right)$
	$\cup \left(\mathbb{V}(a^2 - b^2 - 1) \setminus \mathbb{V}(b^2 + 1, a)\right)$
	$\cup \mathbb{V}(a^2 + b^2 - 1))$
$[y_3, y_2, x_3, x_2]$	$\mathbb{V}(a) \setminus \mathbb{V}(b^2 + 1, a)$
	$\cup \mathbb{V}(a^2 + b^2 - 1) \setminus (\mathbb{V}(b, a - 1) \cup \mathbb{V}(b, a + 1))$
	$\cup \left(\mathbb{V}(a^2 - b^2 - 1) \setminus \left(\mathbb{V}(b^2 + 1, a) \cup \mathbb{V}(b, a - 1) \cup \mathbb{V}(b, a + 1)\right)\right)$
$[y_2, x_3^2, x_2]$	$\mathbb{V}(b, a+1)$
$[y_2, x_3, x_2^2]$	$\mathbb{V}(b, a-1)$

The description of the parameter set (over the reals) for which the theorem is potentially true and no degenerate can be phrased as follows:

> 1) a = 02)  $a^2 + b^2 = 1$  except the points (1,0) and (-1,0) 3)  $a^2 - b^2 = 1$  except the points (1,0) and (-1,0)

This set is represented in Figure 7. and corresponds to



Figure 7: Solutions of Example 6

- 1) The given triangle is itself isosceles (a = 0);
- 2) The given triangle is rectangular at vertex C (with vertices A(-1,0), B(1,0) and the vertex C(a,b) inscribed in the circle  $a^2 + b^2 = 1$ ,
- 3) The given triangle has vertices A(-1,0), B(1,0) and vertex C(a,b) lies on the hyperbola  $a^2 b^2 = 1$ .

Solution 1) is, perhaps, not surprising. Solution 2) corresponds to rectangular triangles for which the orthic triangle reduces to a line, that can be considered a degenerate isosceles triangle. But solution 3) is a nice novelty: it exists a one parameter family of non-isosceles triangles having isosceles orthic triangles.

The remaining two cases in the *MCCGS* output with lpp =  $[y_2, x_3^2, x_2]$  and lpp =  $[y_2, x_3, x_2^2]$  represent degenerate triangles without geometric interest (namely C = A and C = B).

Thus, after performing an automatic proving procedure for the new hypotheses, we can formulate the following theorem:

**Theorem 7.** Given a triangle with vertices A(-1,0), B(1,0) and C(a,b), its orthic triangle will be isosceles if and only if vertex C lies either on the line a = 0 (and then the given triangle is itself isosceles) or in the circle  $a^2 + b^2 = 1$  (and then it is rectangular) or in the hyperbola  $a^2 - b^2 = 1$ .

#### Example 8. [Skaters]

Our final example is taken from the pastimes section of the French journal *Le Monde*, published on the printed edition of Jan. 8, 2007. This example is there attributed to E. Busser and G. Cohen. We think it is nice from *Le Monde* to include the proof of a theorem as a pastime. Actually, the statement to be proved was presented as arising from a more down-to-earth situation: two ice-skaters are moving forming two intersecting circles, at same speed and with the same sense of rotation. They both depart from one of the points



Figure 8: Skaters problem

of intersection of the two circles. Then the journal asked to show that the two skaters were always aligned with the other point of intersection (where some young lady, both skaters were interested at, was placed...).

Let us translate this problem into a theorem discovering question, as follows.

We will consider two circles with centers at P(a, 1) and Q(-b, 1) and radius  $r_1^2 = a^2 + 1$ and  $r_2^2 = b^2 + 1$ , as shown in Figure 8, intersecting at points O(0, 0) and M(0, 2). Consider generic points –the skaters–  $A(x_1, y_1)$  and  $B(x_2, y_2)$  on the respective circles. Point A will be parametrized by the oriented angle v = OPA and, correspondingly, point B will describe the oriented angle w = OQB. Therefore we can say that angle zero corresponds to the departing location of both skaters, namely, point O.

We claim that, for whatever position of points A, B, the points A, M, B are aligned, which is obviously false in general. But we want to determine if there is a relation between the two oriented angles making this statement to hold true. Denote  $c_v, s_v, c_w, s_w$  the cosine and sine of the angles v and w. It is easy to establish the basic hypotheses, using scalar products:

$$H_1 = [(x_1 - a)^2 + (y_1 - 1)^2 - a^2 - 1, (x_2 + b)^2 + (y_2 - 1)^2 - b^2 - 1, a(x_1 - a) + (y_1 - 1) + (a^2 + 1)c_v, -b(x_2 + b) + (y_2 - 1) + (1 + b^2)c_w$$

Now, as the angles are to be taken oriented (because we assume the skaters tare moving on the corresponding circle in the same sense), we need to add the vectorial products involving also the sine to determine exactly the angles and not only their cosines. So we add the hypotheses:

$$H_2 = [a(y_1 - 1) - (x_1 - a) + (a^2 + 1)s_v, -b(y_2 - 1) - (x_2 + b) + (b^2 + 1)s_w]$$

The thesis is, clearly:

$$T = x_1 y_2 - 2x_1 - x_2 y_1 + 2x_2$$

The radii of the circles are

$$r_1^2 = a^2 + 1$$
 and  $r_2^2 = b^2 + 1$ 

and for  $r_1 \neq 0$  and  $r_2 \neq 0$  we have

$$c_{v_0} = \cos v_0 = \cos \widehat{OPM} = \frac{a^2 - 1}{a^2 + 1}, \quad s_{v_0} = \sin v_0 = \sin \widehat{OPM} = \frac{-2a}{a^2 + 1},$$
$$c_{w_0} = \cos w_0 = \cos \widehat{OQM} = \frac{b^2 - 1}{b^2 + 1}, \quad s_{w_0} = \sin w_0 = \sin \widehat{OQM} = \frac{2b}{b^2 + 1}.$$

We want to take a, b and the angles v and w-in terms of the sines and cosines- as parameters. So we must introduce the constraints on the sine and cosine parameters. Moreover, we notice there are also some obvious degenerate situations, namely  $r_1 = 0$ ,  $r_2 = 0$  and a + b = 0, corresponding to null radii or coincident circles, and we want to avoid them.

Currently, *MCCGS* allows us to introduce all these constraints in order to discuss the parametric system. The call is now

$$mccgs(H_1 \cup H_2 \cup T, \ lex(x_1, y_1, x_2, y_2), \ lex(a, b, s_v, c_v, s_w, c_w), \\ null = [c_v^2 + s_v^2 - 1, c_w^2 + s_w^2 - 1], \ notnull = \{a^2 + 1, b^2 + 1, a + b\}).$$

including the constraints on the parameters and eluding degenerate situations as options for *MCCGS*.

The result is that MCCGS outputs only 2 cases. The first one has basis [1], showing that, in general, there is no solution to our query. The second one has  $lpp = [y_2, x_2, y_1, x_1]$  determining in a unique form the points A and B for the given values of the parameters. The associated basis is

$$[y_2 + c_w - bs_w - 1, x_2 - bc_w - s_w + b, y_1 + c_v + as_v - 1, x_1 + ac_v - s_v - a]$$

with parameter conditions that can be expressed as the union of three irreducible varieties:

$$V_1 = \mathbb{V}(c_w^2 + s_w^2 - 1, c_v - c_w, s_v - s_w)$$
  

$$V_2 = \mathbb{V}(c_w^2 + s_w^2 - 1, c_v^2 + s_v^2 - 1, s_w + bc_w - b, bs_w - c_w - 1)$$
  

$$V_3 = \mathbb{V}(c_w^2 + s_w^2 - 1, c_v^2 + s_v^2 - 1, -s_v + ac_v - a, as_v + c_v + 1)$$

The interpretation is easy:  $V_1$  corresponds to arbitrary a, b, w, plus the essential condition v = w, which is the interesting case, stating that our conjecture requires (and it is easy to show that this condition is sufficient) that both skaters keep moving with the same angular speed.

 $V_2$  corresponds to  $s_w = s_{w_0}$ ,  $c_w = c_{w_0}$  and a, b, v free, thus B = M and A can take any position.

 $V_3$  is analogous to  $V_2$ , and corresponds to placing A = M and B anywhere.

So we can summarize the above discussion in the following

**Theorem 9.** Given two non coincident circles of non-null radii and centers P and Q, intersecting at two points O and M, let us consider points A, B on each of the circles. Then the three points A, M, B are aligned if and only if the oriented angles  $\overrightarrow{OPA}$  and  $\overrightarrow{OQB}$  are equal or A or B or both coincide with M.

## 6 Conclusion

We have briefly introduced the principles of automatic discovery and also the ideas –in the context of comprehensive Gröbner basis– for discussing polynomial systems with parameters, via the new *MCCGS* algorithm. Then we have shown how natural is to merge both concepts, since the parameter discussion can be interpreted as yielding, in particular, the projection of the system solution set over the parameter space; and since the conditions for discovery can be obtained by the elimination of the dependent variables over the ideal of hypotheses and thesis.

We have exemplified this approach through a collection of non-trivial examples (performed by running the current Maple implementation of MCCGS, see [MaMo06], over a laptop, without special time – a few seconds– or memory requirements), showing that in all cases, the MCCGS output is very suitable to providing geometric insight, allowing the actual discovery of interesting and new? theorems (and pastimes!).

# References

- [BDR] C. Beltrán, G. Dalzotto, T. Recio. The moment of truth in automatic theorem proving in elementary geometry. Proceedings ADG 2006 (extended abstracts). F. Botana, E. Roanes-Lozano, eds. Universidad de Vigo, 2006.
- [BR05] F. Botana, T. Recio. Towards solving the dynamic geometry bottleneck via a symbolic approach. Proceedings ADG (Automatic Deduction in Geometry) 2004. Springer. Lec. Not. Artificial Intelligence LNAI 3763, pp. 92-111, 2006.
- [CLLW] X. F. Chen, Peng Li, Long Lin, D. K. Wang. Proving geometric theorems by partitioned-parametric Gröbner bases. Proceedings ADG (Automatic Deduction in Geometry) 2004. Springer. Lec. Not. Artificial Intelligence LNAI 3763, pp. 34-44, 2005.
- [CW] X. F. Chen, D. K. Wang The projection of a quasivariety and its application on goemetry theorem proving and formula deduction. Proceedings ADG (Automatic Deduction in Geometry) 2002. Springer. Lec. Not. Artificial Intelligence LNAI 2930, pp. 21-30, 2004.
- [Ch84] S.-C. Chou. Proving Elementary Geometry Theorems Using Wu's Algorithm Contemporary Mathematics, vol.29, 243-286, Automated Theorem Proving: After 25 Years, American Mathematical Society, Providence, Rhode Island, 1984.
- [Ch87] S.C. Chou. A Method for Mechanical Derivation of Formulas in Elementary Geometry. Journal of Automated Reasoning, 3, 291–299, 1987.
- [Ch88] S.-C. Chou. Mechanical Geometry Theorem Proving. Mathematics and its Applications, D. Reidel Publ. Comp. (1988).
- [ChG90] S.-C. Chou, X. -S. Gao. Methods for Mechanical Geometry Formula Deriving Proceedings of International Symposium on Symbolic and Algebraic Computation, 265-270, ACM Press, 1990.

- [CNR99] A. Capani, G. Niesi, L. Robbiano. CoCoA, a System for Doing Computations in Commutative Algebra. The version 4.6 is available at the web site http://cocoa.dima.unige.it.
- [DR] G. Dalzotto, T. Recio On protocols for the automated discovery of theorems in elementary geometry. submitted J. Automated Reasoning
- [K86] D. Kapur. Using Gröbner basis to reason about geometry problems. J. Symbolic Computation, 2(4), 399-408, 1986.
- [K89] D. Kapur. Wu's method and its application to perspective viewing. Geometric Reasoning. D. Kapur and J.L. Mundy, eds. The MIT press, Cambridge, Ma., 1989.
- [K95] D. Kapur. An approach to solving systems of parametric polynomial equations. Principles and Practice of Constraint Programming. Saraswat and Van Hentenryck, eds. MIT Press, Cambridge, Ma., 1995.
- [KR00] M. Kreuzer, L. Robbiano. Computational Commutative Algebra 1. Springer Verlag (2000).
- [Ko] W. Koepf Gröbner bases and triangles The International Journal of Computer Algebra in Mathematics Education. Vol. 4, No.4, pp. 371-386, 1998.
- [MaMo05] M. Manubens, A. Montes. Improving DISPGB Algorithm Using the Discriminant Ideal. Jour. Symb. Comp. 41 (2006) 1245–1263.
- [MaMo06] M. Manubens, A. Montes. Minimal Canonical Comprehensive Groebner System. arXiv: math.AC/0611948. (2006).
- [Mo02] A. Montes. New Algorithm for Discussing Gröbner Bases with Parameters. Jour. Symb. Comp. 33:1-2 (2002) 183–208.
- [Mo06] A. Montes. About the canonical discussion of polynomial systems with parameters. arXiv: math.AC/0601674. (2006).
- [R98] T. Recio. Cálculo simbólico y geométrico. Editorial Síntesis. Madrid. 1998.
- [RV99] T. Recio, M. Pilar Vélez. Automatic Discovery of Theorems in Elementary Geometry. J. Automat. Reason. 23: 63–82 (1999).
- [RB] T. Recio, F. Botana. Where the truth lies (in automatic theorem proving in elementary geometry). Proceedings ICCSA (International Conference on Computational Science and its Applications) 2004. Springer. Lec. Not. Com. Sci. 3044, pp. 761-771, 2004.
- [Ri] P. Richard, Raisonnement et stratégies de preuve dans l'enseignement des mathématiques. Peter Lang Editorial, Berne, 2004.
- [Wa98] D. Wang. Gröbner Bases Applied to Geometric Theorem Proving and Discovering. Gröbner Bases and Applications, B. Buchberger and F. Winkler eds. London Mathematical Society Lecture Notes Series Vol 251: 281–301, Cambridge University Press (1998).

- [Weis92] V. Weispfenning. Comprehensive Grobner bases. Journal of Symbolic Computation, v.14 n.1, p.1-29, July 1992.
- [Wib06] M. Wibmer. Gröbner bases for families of affine or projective schemes. arXiv: math.AC/0608019. (2006)

Antonio Montes Departament de Matemàtica Aplicada 2, Universitat Politècnica de Catalunya, Spain. e-mail: antonio.montes@upc.edu http://www-ma2.upc.edu/~montes

Tomas Recio Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Spain. e-mail: tomas.recio@unican.es http://www.recio.tk