Detecting unnecessary assumptions of elementary geometry problems by CAS

Yosuke Sato [1] *and Ryoya Fukasaku* [2]
ysato@rs.kagu.tus.ac.jp [1]   1414704@ed.tus.ac.jp [2]
Department of Mathematical Information Science
Tokyo University of Science
162-8601, Tokyo, Japan

Abstract

A geometry problem given for highschool students in a place such as a competition of International Mathematical Olympiad or an entrance examination for a university often contains unnecessary assumptions. Such assumptions sometimes undermine the quality of the problem. In this paper we introduce a method to check whether a given assumption is essentially necessary for solving the problem. It also computes necessary and sufficient assumptions for getting a solution. Our method uses three tools of computer algebra, Gröbner bases computation, quantifier elimination over complex numbers and quantifier elimination over real numbers. Anyone can use our method with a minimum basic knowledge of computer algebra using any CAS with those tools.

1 Introduction

A mathematical problem given for highschool students in a worldwide high level competition such as International Mathematical Olympiad should be sophisticated. Unfortunately, there are many geometry problems of International Mathematical Olympiad which have unnecessary assumptions for getting their solutions. For example, in the following problem the triangle ABC is assumed to be acute-angled, but we do not need it for proving the consequence.

Problem 4 (International Mathematical Olympiad 2013)

Let ABC be an acute-angled triangle with orthocentre H, and let W be a point on the side BC, lying strictly between B and C. The points M and N are the feet of the altitudes from B and C, respectively. Denote by ω_1 the circumcircle of BWN, and let X be the point on ω_1 such that WX is a diameter of ω_1. Analogously, denote by ω_2 the circumcircle of CWM, and let Y be the point on ω_2 such that WY is a diameter of ω_2. Prove that X, Y and H are collinear.

The authors believe that a good problem should have necessary and sufficient assumptions for solving it. In this paper, with a focus on elementary geometry problems, we introduce a method using CAS to check whether an assumption of a problem is essentially necessary for solving it.

Any formula of elementary geometry can be represented as a first order formula constructed from polynomial equations and inequalities. The structure of $\mathbb{R}$ is a complete model for Tarski geometry. Since quantifier elimination of such a first order formula (called *real QE* in this paper) is computable [9, 3], for any given problem of elementary geometry, theoretically we can compute necessary and sufficient assumptions for getting the solution. For the above problem, assigning the points A, B and C with coordinates $(0,0),(1,0)$ and (c_1,c_2) respectively, we can obtain a necessary and sufficient assumption of the triangle ABC by computing a quantifier free formula of c_1,c_2 which is equivalent to the following formula. (See Section 3.2 for the details.)

$$\forall x_1,x_2,y_1,y_2,m_1,m_2,h_2,w\in\mathbb{R}(\\ 0<w<1\wedge F_1=0\wedge F_2=0\wedge F_3=0\wedge F_4=0\wedge F_5=0\wedge F_6=0\wedge F_7=0\Rightarrow P=0).$$

Unfortunately, however, we cannot eliminate the quantifiers in a feasible length of time by any of the existing software of real QE such as the Mathematica packages Reduce and Resolve ([6]), a real QE program in Regular Chains Package of Maple ([1]), QEPCAD ([7]), real QE programs rlqe and rlhqe of redlog ([8]) except for the recent real QE implementation introduced in [5]. Computations of real QE are very heavy in general. There are many geometry problems given in the past competitions of International Mathematical Olympiad which contain unnecessary assumptions but we cannot detect it by any of existing software of real QE.

There are four hierarchies of elementary geometry, affine geometry, metric geometry, Hilbert geometry and Tarski geometry. In order to present a formula algebraically, we need only equation $=$ and disequation $\neq$ in affine and metric geometry, but we essentially need inequality $>$ in Hilbert and Tarski geometry. (More detailed descriptions may be found in a article such as [2].) The structure of $\mathbb{C}$is a complete model for metric geometry, hence we can deal with any problem of metric geometry by the computation of the underlying polynomial ideal. For a problem of Hilbert geometry, we cannot deal with it by the computation of the underlying polynomial ideal only but we can reduce the whole problem to a much smaller problem using a decomposition of the ideal. The reduced problem can be easily handled by computation of real QE. Most (probably all) geometry problems given in the past International Mathematical Olympiad are at most in Hilbert geometry.

In this paper, we introduce a method to compute necessary and sufficient assumptions for a given problem. Our method consists of three devises. The first one applies to a problem of metric geometry which has no hidden non-degenerate assumptions. It uses only computation of a Gröbner basis. The second one applies to a problem of metric geometry which has hidden non-degenerate assumptions. It uses computation of complex QE, i.e. QE over $\mathbb{C}$. The third one applies to a problem of Hilbert geometry. It computes a decomposition of the underlying ideal by factorization of polynomials then reduce the whole problem to simpler subproblems of real QE. Anyone can use our method with a minimum basic knowledge of computer algebra using any CAS with those tools.

The paper is organized as follows. In section 2, we give a minimum description of mathematical background to understand our paper. Since there is no standard representation form of elementary geometry problems, it is not easy to describe our method as concrete algorithms. Using typical examples of geometry problems given in the past International Mathematical Olympiad, we describe our method. Our method is easily applied to any other problem. In section 3, we treat two types of problems of metric geometry, one is a problem with no hidden non-degenerate assumptions and the other is a problem with hidden non-degenerate assumptions. In section 4, we treat a problem of Hilbert geometry.

2 Preliminaries

$\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ denote the field of rational numbers, real numbers and complex numbers respectively. For a big letter such as X, $\bar{X}$ denotes some variables $X_1, \ldots, X_n$. $T(\bar{X})$ denotes a set of terms in $\bar{X}$. An admissible term order of $T(\bar{X}, \bar{Y})$ such that any variable X_i is greater than any term of $T(\bar{Y})$ is denoted by $\bar{X} \gg \bar{Y}$. For an ideal $I \subset \mathbb{Q}[\bar{X}]$, $\mathbb{V}(I)$ denotes its variety in $\mathbb{C}$.

2.1 Gröbner Bases

The following fundamental properties of Gröbner bases, found in most standard text books of Gröbner bases, play important roles in this paper.

Theorem 1 *Let I be an ideal in a polynomial ring $\mathbb{Q}[\bar{X}]$. For any admissible term order of $T(\bar{X})$, $\mathbb{V}(I) = \emptyset$ if and only if the reduced Gröbner basis of I is equal to $\{1\}$.*

Corollary 2 *For polynomials $f_1(\bar{X}), \ldots, f_l(\bar{X}), h(\bar{X}), g(\bar{X})$ in $\mathbb{Q}[\bar{X}]$,
$\forall \bar{a} \in \mathbb{C}^n (f_1(\bar{a}) = 0 \wedge \cdots \wedge f_l(\bar{a}) = 0 \wedge h(\bar{a}) \neq 0 \Rightarrow g(\bar{a}) = 0)$ holds if and only if the reduced Gröbner basis of the ideal $\langle f_1, \ldots, f_l, hgY - 1\rangle$ in $\mathbb{Q}[\bar{X}, Y]$ is equal to $\{1\}$ for any admissible term order of $\bar{X}, Y$.*

Theorem 3 *Let I be an ideal in a polynomial ring $\mathbb{Q}[\bar{X}, \bar{Y}]$. Let G be a Gröbner basis of I w.r.t. a term order such that $\bar{X} \gg \bar{Y}$, then $G \cap \mathbb{Q}[\bar{Y}]$ is a Gröbner basis of the elimination ideal $I \cap \mathbb{Q}[\bar{Y}]$ w.r.t. the same term order.*

2.2 Complex QE

Let ϕ be a first-order formula with atomic formulas of polynomial equations over $\mathbb{Q}$. QE of ϕ over $\mathbb{C}$ is to obtain its equivalent quantifier free formula in the structure of $\mathbb{C}$. Using a prenex normal form of a given formula, we can reduce any QE problem to a QE problem of the following basic formula with polynomials $f_1, \ldots, f_s, g$ of $\mathbb{Q}[\bar{X}, \bar{Y}]$:

$$\exists \bar{X} \in \mathbb{C}^n (f_1(\bar{X}, \bar{Y}) = 0 \wedge \cdots \wedge f_s(\bar{X}, \bar{Y}) = 0 \wedge g(\bar{X}, \bar{Y}) \neq 0).$$

The problem is computable and plays an important role in this paper. There exist several implementations of complex QE. Mathematica package Reduce and Resolve contain complex QE implementation, Maple package Projection can handle the above basic formula. Besides these programs, we can easily obtain the quantifier free formula of the basic formula by computation of a comprehensive Gröbner systems (CGSs). Recent implementation of complex QE reported in [4] which is based on the computation of a CGS compute the simplest quantifier free formula in most cases.

2.3 Real QE

Let ϕ be a first-order formula with atomic formulas of polynomial equations and inequalities over $\mathbb{Q}$. QE of ϕ over $\mathbb{R}$ is to obtain its equivalent quantifier free formula in the structure of $\mathbb{R}$. Using a prenex normal form of a given formula, we can reduce any QE problem to a QE problem of the following basic formula with polynomials $f_1, \ldots, f_s, g_1, \ldots, g_t$ of $\mathbb{Q}[\bar{Y}, \bar{X}]$:

$$\exists \bar{X} \in \mathbb{R}^n (f_1(\bar{X}, \bar{Y}) = 0 \wedge \cdots \wedge f_s(\bar{X}, \bar{Y}) = 0 \wedge g_1(\bar{X}, \bar{Y}) > 0 \wedge \cdots \wedge g_t(\bar{X}, \bar{Y}) > 0).$$

The problem is also computable and plays an important role in this paper. We call such QE *real QE* in this paper. There exist several implementations of real QE, Mathematica package Reduce and Resolve contain real QE implementation, the redlog package of the CAS REDUCE contains two real QE programs rlqe and rlhqe ([8]), an open source real QE program QEPCAD ([7]), the real QE program in Regular Chains package of Maple ([1]) and a recent real QE implementation introduced in [5], etc., are available.

3 Problems of Metric Geometry

In this section, we introduce two methods to check whether a given problem of elementary geometry is a problem of metric geometry. The first one applies to a problem with no hidden non-degenerate assumptions. If a given problem is a problem of metric geometry and it has only obvious non-degenerate assumptions, the first method detects it and finds all unnecessary assumptions for solving it. If a given problem is a problem of metric geometry but has some hidden non-degenerate assumptions, the second one finds all such assumptions together with all unnecessary assumptions for solving it.

3.1 Problems with no hidden non-degenerate assumptions

Consider the following problem. We will show that it is a problem of metric geometry and has an unnecessary assumption for solving it by computation of a Gröbner basis.

Problem 1 (International Mathematical Olympiad 2012)

Given triangle ABC the point J is the centre of the excircle opposite the vertex A. This excircle is tangent to the side BC at M, and to the lines AB and AC at K and L, respectively. The lines LM and BJ meet at F, and the line KM and CJ meet at G. Let S be the point of intersection of the lines AF and BC, and let T be the point of intersection of the lines AG and BC. Prove that M is the midpoint of ST.
(The excircle of ABC opposite the vertex A is the circle that is tangent to the line segment BC, to the ray AB beyond B, and to the ray AC beyond C.)

Let the coordinates of B and C be $(0, 0)$ and $(1, 0)$ respectively w.l.o. generality. Let the coordinates of A and J be (a_1, a_2) and (j_1, j_2) respectively. Since S and T are on the line BC, the coordinates of S and T are $(s_1, 0)$ and $(t_1, 0)$ respectively.

1. We need an obvious non-degenerate assumption $a_2 \neq 0$ in order that the points A, B and C are not collinear.
2. Since JM $\perp$ BC, the coordinate of M is $(j_1, 0)$.
3. Since K is on the line AB such that B is between A and K, the coordinate of K is $(-ka_1, -ka_2)$ for some positive real number k.
4. Since L is on the line AC such that C is between A and L, the coordinate of L is $(1 - l(a_1 - 1), -la_2)$ for some positive real number l.
5. Since F is on the line BJ, the coordinate of F is (fj_1, fj_2) for some real number f.
6. Since G is on the line CJ, the coordinate of G is $(1 + g(j_1 - 1), gj_2)$ for some real number g.
7. The condition that BA $\perp$ KJ (equivalent to KA $\perp$ KJ) is represented by the following equation: $a_1(j_1 - (-ka_1)) + a_2(j_2 - (-ka_2)) = 0$.
8. The condition that CA $\perp$ LJ (equivalent to LA $\perp$ LJ), is represented by the following

equation: $(a_1-1)(j_1-(1-l(a_1-1)))+a_2(j_2-(-la_2))=0$.

9. The condition that G, M and K are collinear is represented by the following equation: $((1+g(j_1-1))-j_1)(-ka_2)-(gj_2)((-ka_1)-j_1)=0$.
10. The condition that F, M and L are collinear is represented by the following equation: $((fj_1)-j_1)(-la_2)-(fj_2)((1-l(a_1-1))-j_1)=0$.
11. The condition that A, G and T are collinear is represented by the following equation: $(a_1-t_1)(gj_2)-a_2((1+g(j_1-1))-t_1)=0$.
12. The condition that A, F and S are collinear is represented by the following equation: $(a_1-s_1)(fj_2)-a_2((fj_1)-s_1)=0$.
13. The condition |KJ| = |MJ| is represented by the following equation: $(j_1-(-ka_1))^2+(j_2-(-ka_2))^2-j_2^2=0$
14. The condition |LJ| = |MJ| is represented by the following equation: $(j_1-(1-l(a_1-1)))^2+(j_2-(-la_2))^2-j_2^2=0$.
15. The condition that M is the midpoint of ST is represented by the following equation: $t_1+s_1-2j_1=0$

Let $F_1=a_1(j_1-(-ka_1))+a_2(j_2-(-ka_2))$, $F_2=(a_1-1)(j_1-(1-l(a_1-1)))+a_2(j_2-(-la_2))$, $F_3=((1+g(j_1-1))-j_1)(-ka_2)-(gj_2)((-ka_1)-j_1)$, $F_4=((fj_1)-j_1)(-la_2)-(fj_2)((1-l(a_1-1))-j_1)$, $F_5=(a_1-t_1)(gj_2)-a_2((1+g(j_1-1))-t_1)$, $F_6=(a_1-s_1)(fj_2)-a_2((fj_1)-s_1)$, $F_7=(j_1-(-ka_1))^2+(j_2-(-ka_2))^2-j_2^2$, $F_8=(j_1-(1-l(a_1-1)))^2+(j_2-(-la_2))^2-j_2^2$, $P=t_1+s_1-2j_1$.

The problem is nothing but proving the following sentence is true:

$$\forall a_1,a_2,k,l,g,f,j_1,j_2,s_1,t_1\in\mathbb{R}(a_2\neq 0\wedge k>0\wedge l>0\wedge F_1=0\wedge F_2=0\wedge F_3=0\wedge F_4=0\wedge F_5=0\wedge F_6=0\wedge F_7=0\wedge F_8=0\Rightarrow P=0).$$

In fact we do not need any assumption on the excenter J. The conclusion $P=0$ holds not only for other 3 excenter but also for the incenter. That is the above sentence is true even if we exclude the conditions $k>0$ and $l>0$.

Unfortunately, some real QE implementations such as Mathematica packages Resolve and Reduce do not terminate in hours for these QE problems. Note that after we remove all conditions containing inequalities from an original problem, if the problem still holds then the original problem is likely to belong to metric geometry. Remember that the structure of $\mathbb{C}$ is a complete model for metric geometry, hence it suffices to check the following sentence is true.

$$\forall a_1,a_2,k,l,g,f,j_1,j_2,s_1,t_1\in\mathbb{C}(a_2\neq 0\wedge F_1=0\wedge F_2=0\wedge F_3=0\wedge F_4=0\wedge F_5=0\wedge F_6=0\wedge F_7=0\wedge F_8=0\Rightarrow P=0).$$

By Corollary 2, the sentence is true if and only if the reduced Gröbner basis of the ideal $\langle F_1,\dots,F_8,a_2PY-1\rangle\subset\mathbb{Q}[a_1,a_2,k,l,g,f,j_1,j_2,s_1,t_1,Y]$ is equal to $\{1\}$.

Gröbner bases computation program on most CAS such as Mathematica, Maple, REDUCE, Singular and Risa/Asir, etc. immediately returns the above output $\{1\}$. It shows that the problem is of metric geometry and the given assumption for the excenter is too severe.

3.2 Problems with hidden non-degenerate assumptions

The problem discussed in the previous subsection has no hidden non-degenerate assumptions. If a problem contains some hidden, i.e. non trivial, non-degenerate assumptions, unless we add all such assumptions we cannot detect unnecessary assumptions. Complex QE is useful for finding all hidden non-degenerate assumptions. Consider the problem given in the introduction.

Let the coordinate of A and B be $(0, 0)$ $(1, 0)$ respectively w.l.o. generality.
Let the coordinate of C, M, X and Y be (c_1, c_2),(m_1, m_2), (x_1, x_2) and (y_1, y_2) respectively.

1. We add the condition $c_2 \neq 0$ in order that the points A, B and C are not collinear.
2. The condition that ABC is an acute-angled triangle is represented by the following formula: $0 < c_1 < 1 \wedge (c_1 - 1/2)^2 + c_2^2 > 1/4$.
3. Since CN $\perp$ AB, the coordinate of N is $(c_1, 0)$.
4. Since H is on the line CN, its coordinate is (c_1, h_2) for some real number h_2.
5. Since W is on the line BC, its coordinate is $(1 + w(c_1 - 1), wc_2)$ for some real number w. Since W is strictly between B and C, we need $0 < w < 1$.
6. The condition that M is on the line AC is represented by the following equation: $m_1c_2 - m_2c_1 = 0$.
7. The condition that BH $\perp$ AC is represented by the following equation: $(c_1 - 1)c_1 + h_2c_2 = 0$.
8. The condition that BM $\perp$ AC, which is equivalent to that H is on the line BM, is represented by the following equation: $(m_1 - 1)c_1 + m_2c_2 = 0$.
9. The condition that WX is a diameter of ω_1 is represented by the following equations:

$$((1 + w(c_1 - 1)) - x_1)^2 + (wc_2 - x_2)^2 = ((1 + w(c_1 - 1)) + x_1 - 2c_1)^2 + (wc_2 + x_2)^2$$
$$= ((1 + w(c_1 - 1)) + x_1 - 2)^2 + (wc_2 + x_2)^2.$$

10. The condition that WY is a diameter of ω_2 is represented by the following equations:

$$((1 + w(c_1 - 1)) - y_1)^2 + (wc_2 - y_2)^2 = ((1 + w(c_1 - 1)) + y_1 - 2mc_1)^2 + (wc_2 + y_2 - 2mc_2)^2$$
$$= ((1 + w(c_1 - 1)) + y_1 - 2c_1)^2 + (wc_2 + y_2 - 2c_2)^2.$$

11. The condition that X, Y and H are collinear is represented by the following equation: $(y_1 - c_1)(y_2 - x_2) - (y_2 - h_2)(y_1 - x_1) = 0$.

Let $F_1 = m_1c_2 - m_2c_1$, $F_2 = (c_1 - 1)c_1 + h_2c_2$, $F_3 = (m_1 - 1)c_1 + m_2c_2$, $F_4 = ((1 + w(c_1 - 1)) - x_1)^2 + (wc_2 - x_2)^2 - (((1 + w(c_1 - 1)) + x_1 - 2c_1)^2 + (wc_2 + x_2)^2)$, $F_5 = ((1 + w(c_1 - 1)) - x_1)^2 + (wc_2 - x_2)^2 - (((1 + w(c_1 - 1)) + x_1 - 2)^2 + (wc_2 + x_2)^2)$, $F_6 = ((1 + w(c_1 - 1)) - y_1)^2 + (wc_2 - y_2)^2 - (((1 + w(c_1 - 1)) + y_1 - 2mc_1)^2 + (wc_2 + y_2 - 2mc_2)^2)$, $F_7 = ((1 + w(c_1 - 1)) - y_1)^2 + (wc_2 - y_2)^2 - (((1 + w(c_1 - 1)) + y_1 - 2c_1)^2 + (wc_2 + y_2 - 2c_2)^2)$, $P = (y_1 - c_1)(y_2 - x_2) - (y_2 - h_2)(y_1 - x_1)$.

The problem is nothing but proving the following sentence is true:

$$\forall x_1, x_2, y_1, y_2, m_1, m_2, h_2, w, c_1, c_2 \in \mathbb{R}(c_2 \neq 0 \wedge 0 < c_1 < 1 \wedge (c_1 - 1/2)^2 + c_2^2 > 1/4 \wedge$$
$$0 < w < 1 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge F_6 = 0 \wedge F_7 = 0 \Rightarrow P = 0).$$

Unfortunately, any of the real QE program, Resolve or Reduce in Mathematica, the real QE program of RegularChains package in Maple, the real QE software QEPCAD, the real QE programs rlqe and rlhqe in redlog is unable to handle it. They cannot either compute a necessary and sufficient assumption of c_1 and c_2 for getting the solution as is described in the introduction.

Note that the problem contains an obvious non-degenerate assumption $c_2 \neq 0$ in order that the points A, B and C are not collinear. It also contains obvious non-degenerate assumptions $w \neq 0, 1$ in order that BWN and CWM have their circumcircle. Using only these conditions, we can see the problem is actually a problem of metric geometry together with getting all hidden non-degenerate assumptions by applying complex QE to the following formula:

$$\forall x_1, x_2, y_1, y_2, m_1, m_2, h_2, w \in \mathbb{C}(w \neq 0 \wedge w \neq 1 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge F_6 = 0 \wedge F_7 = 0 \Rightarrow P = 0).$$

The equivalent quantifier free formula is the following:

$$(c_1 = 1 \wedge c_2^2 + 1 = 0) \vee (c_1 = 0 \wedge c_2 = 0) \vee (c_1 - 1)((c_1 - 1/2)^2 + c_2^2 - 1/4) \neq 0.$$

The first formula $c_1 = 1 \wedge c_2^2 + 1 = 0$ is impossible for real values and the second formula $c_1 = 0 \wedge c_2 = 0$ is also impossible under the condition $c_2 \neq 0$. Hence, the following formula is true:

$$\forall x_1, x_2, y_1, y_2, m_1, m_2, h_2, w, c_1, c_2 \in \mathbb{R}(c_1 \neq 1 \wedge c_2 \neq 0 \wedge (c_1 - 1/2)^2 + c_2^2 \neq 1/4 \wedge w \neq 0 \wedge w \neq 1 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge F_6 = 0 \wedge F_7 = 0 \Rightarrow P = 0).$$

Note that the condition $c_1 = 1$ implies N=B and the condition $(c_1 - 1/2)^2 + c_2^2 = 1/4$ implies M=C, both are degenerate cases. Note also that $c_1 = 1 \Leftrightarrow$ CB⊥BA and $(c_1 - 1/2)^2 + c_2^2 = 1/4 \Leftrightarrow$ AC⊥CB. Hence, necessary and sufficient conditions for the conclusion are ∠ABC$\neq \pi/2$, ∠ACB$\neq \pi/2$ and W≠B,W≠C. The important point is that we can automatically obtain them by the computation of complex QE and detect that the problem belongs to metric geometry. Our implementation of complex QE ([4]) computes the above quantifier free formula within 1 second in a standard laptop computer. The Maple package Projection also computes it within a few seconds.

4 Problems of Hilbert Geometry

Let O be the origin $(0, 0)$. Given four points T= (t_1, t_2), U= (u_1, u_2), V= (v_1, v_2) and W= (w_1, w_2) besides the origin. Then, the relation ∠TOU= ∠VOW is represented by the equation $\frac{t_1u_1+t_2u_2}{\sqrt{(t_1^2+t_2^2)(u_1^2+u_2^2)}} = \frac{v_1w_1+v_2w_2}{\sqrt{(v_1^2+v_2^2)(w_1^2+w_2^2)}}$. But this is not a polynomial equation.
In order to represent it by a polynomial equation, we need to use inequality as follows:

$$\angle\text{TOU}= \angle\text{VOW} \Leftrightarrow (t_1u_1 + t_2u_2)^2(v_1^2 + v_2^2)(w_1^2 + w_2^2) = (v_1w_1 + v_2w_2)^2(t_1^2 + t_2^2)(u_1^2 + u_2^2) \wedge (t_1u_1 + t_2u_2)(v_1w_1 + v_2w_2) \geq 0.$$

Consider the following problem. It contains an equation between two angles. Such a problem is likely to belong to Hibert geometry, since we essentially need an inequality as described above.

Problem 4 (International Mathematical Olympiad 2014)

Points P and Q lie on side BC of acute-angled triangle ABC so that ∠ PAB = ∠ BCA and ∠ CAQ = ∠ ABC. Points M and N lie on lines AP and AQ, respectively, such that P is the midpoint of AM, and Q is the midpoint of AN. Prove that lines BM and CN intersect on the circumcircle of triangle ABC.

Let the coordinates of A, B and C be (a_1, a_2), $(0, 0)$ and $(1, 0)$ respectively w.l.o. generality. Let the coordinates of P and Q be $(p_1, 0)$ and $(q_1, 0)$. Let the intersection point of BM and CN be E and its coordinate (e_1, e_2). Let D be the circumcenter of the circumcircle of triangle ABC.

1. We add the condition $a_2 \neq 0$ in order that the points A, B and C are not collinear.
2. The coordinates of M and N are $(-a_1 + 2q_1, -a_2)$ and $(-a_1 + 2p_1, -a_2)$ respectively by the assumption.

3. The coordinate of D is $(1/2, d_2)$ for some real number d_2 by the assumption DB=DC.
4. The condition that ABC is an acute-angled triangle is represented by the following formula:
$$0 < a_1 < 1 \wedge (a_1 - 1/2)^2 + a_2^2 > 1/4.$$
5. The condition that D is the circumcircle of triangle ABC, which is equivalent to $|\mathrm{AD}| = |\mathrm{BD}| = |\mathrm{CD}|$, is represented by the following equation:
$$(a_1 - 1/2)^2 + (a_2 - d_2)^2 - (d_2^2 + 1/4) = 0.$$
6. The condition that the points E is on the line NC is represented by the following equation:
$$e_2(-a_1 + 2q_1 - 1) + a_2(e_1 - 1) = 0.$$
7. The condition that the points E is on the line MB is represented by the following equation:
$$e_2(-a_1 + 2p_1) + a_2e_1 = 0.$$
8. The condition that ∠PAB= ∠BCA is represented by the following equation and inequality:
$$(-a_1(p_1 - a_1) + a_2^2)^2((a_1 - 1)^2 + a_2^2) - (1 - a_1)^2(a_1^2 + a_2^2)((p_1 - a_1)^2 + a_2^2) = 0$$
$$\wedge(-a_1(p_1 - a_1) + a_2^2)(1 - a_1) \geq 0.$$
9. The condition that ∠CAQ= ∠ABC is represented by the following equation and inequality:
$$((1 - a1)(q_1 - a_1) + a_2^2)^2(a_1^2 + a_2^2) - a_1^2((1 - a_1)^2 + a_2^2)((q_1 - a_1)^2 + a_2^2) = 0$$
$$\wedge((1 - a_1)(q_1 - a_1) + a_2^2)a_1 \geq 0.$$
10. The condition that the points P and Q lie on side BC is represented by the inequalities:
$$1 \geq p1 \geq 0, 1 \geq q1 \geq 0.$$
11. The conclusion which is equivalent to $|\mathrm{ED}| = |\mathrm{DB}|$ is represented by the equation:
$$(e_1 - 1/2)^2 + (e_2 - d_2)^2 - (d_2^2 + 1/4) = 0.$$

Let $F_1 = (a_1 - 1/2)^2 + (a_2 - d_2)^2 - (d_2^2 + 1/4), F_2 = e_2(-a_1 + 2q_1 - 1) + a_2(e_1 - 1), F_3 = e_2(-a_1 + 2p_1) + a_2e_1, F_4 = (-a_1(p_1 - a_1) + a_2^2)^2((a_1 - 1)^2 + a_2^2) - (1 - a_1)^2(a_1^2 + a_2^2)((p_1 - a_1)^2 + a_2^2), F_5 = ((1 - a1)(q_1 - a_1) + a_2^2)^2(a_1^2 + a_2^2) - a_1^2((1 - a_1)^2 + a_2^2)((q_1 - a_1)^2 + a_2^2), G_1 = (-a_1(p_1 - a_1) + a_2^2)(1 - a_1), G_2 = ((1 - a_1)(q_1 - a_1) + a_2^2)a_1, P = (e_1 - 1/2)^2 + (e_2 - d_2)^2 - (d_2^2 + 1/4)$.

The problem is nothing but proving the following sentence is true:
$$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{R}(a_2 \neq 0 \wedge 1 > a_1 > 0 \wedge (a_1 - 1/2)^2 + a_2^2 > 1/4 \wedge 1 \geq p_1 \geq 0 \wedge 1 \geq q_1 \geq 0 \wedge$$
$$F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge G_1 \geq 0 \wedge G_2 \geq 0 \Rightarrow P = 0).$$

In fact, we can see the following sentence is true:
$$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{R}(a_2 \neq 0 \wedge p_1 \geq 0 \wedge 1 \geq q_1 \wedge$$
$$F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge G_1 \geq 0 \wedge G_2 \geq 0 \Rightarrow P = 0).$$

Unfortunately, most real QE programs are unable to handle both of them except for the real QE program of RegularChains package in Maple and the recent real QE implementation of [5].

Using the methods introduced in the previous section, we can check both of the following sentence and formula is false:

$$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \Rightarrow P = 0),$$
$$\forall a_1, a_2, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \Rightarrow P = 0).$$

As we predicted, the problem does not belong to metric geometry. Using computation of a Gröbner basis we can check that it is a problem of Hilbert geometry and obtain necessary and sufficient conditions for getting the conclusion.

For a problem of a Hilbert geometry, an essential need of an inequality arises for fixing an angle. For an angle $\theta(\pi > \theta > 0)$, $\pi/2 \geq \theta$ if and only if $\cos\theta \geq 0$. When we know the value $\cos\theta$, even though we cannot fix an angle θ without using an inequality, we can narrow its possible values to only two candidates. In our problem, we can narrow the range of the coordinate of the point P or Q to two possible candidates in terms of a_1, a_2. Since there are four possible combinations, we can decompose the underlying ideal $I = \langle a_2 v - 1, F_1, \ldots, F_5 \rangle \subset \mathbb{Q}[v, a_1, a_2, p_1, q_1, d_2, e_1, e_2]$ to $I_1 \cap I_2 \cap I_3 \cap I_4$. In each component I_i, the coordinates of P and Q are uniquely determined. By computing a Gröbner basis of I w.r.t. a term order such that $v, q_1, d_2, e_1, e_2 \gg a_1, a_2, p_1$ we obtain the elimination ideal $I \cap \mathbb{Q}[a_1, a_2, p_1] = \langle a_1^4 + 2a_1^2a_2^2 + a_2^4 - 2a_1^3p_1 - 2a_1a_2^2p_1 - p_1^2 + 2a_1p_1^2 \rangle$, a term order such that $v, p_1, d_2, e_1, e_2 \gg a_1, a_2, q_1$ yields the elimination ideal $I \cap \mathbb{Q}[a_1, a_2, q_1]$ $= \langle -2a_1^3 + a_1^4 - 2a_1a_2^2 + 2a_1^2a_2^2 + a_2^4 - 2a_1q_1 + 6a_1^2q_1 - 2a_1^3q_1 + 2a_2^2q_1 - 2a_1a_2^2q_1 + q_1^2 - 2a_1q_1^2 \rangle$. Factorizing two generator polynomials we get the following decomposition:

$I = (I + \langle a_1^2 + a_2^2 - p_1, -2a_1 + a_1^2 + a_2^2 + q_1 \rangle) \cap (I + \langle a_1^2 + a_2^2 - p_1, a_1^2 + a_2^2 + q_1 - 2a_1q_1 \rangle) \cap (I + \langle a_1^2 + a_2^2 + p_1 - 2a_1p_1, -2a_1 + a_1^2 + a_2^2 + q_1 \rangle) \cap (I + \langle a_1^2 + a_2^2 + p_1 - 2a_1p_1, a_1^2 + a_2^2 + q_1 - 2a_1q_1 \rangle)$

By the method given in Section 3.1, we can check only the first one of the following four sentences is true:

$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge,$
$a_1^2 + a_2^2 - p_1 = 0 \wedge -2a_1 + a_1^2 + a_2^2 + q_1 = 0 \Rightarrow P = 0),$

$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge,$
$a_1^2 + a_2^2 - p_1 = 0 \wedge a_1^2 + a_2^2 + q_1 - 2a_1q_1 = 0 \Rightarrow P = 0),$

$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge,$
$a_1^2 + a_2^2 + p_1 - 2a_1p_1 = 0 \wedge -2a_1 + a_1^2 + a_2^2 + q_1 = 0 \Rightarrow P = 0),$

$\forall a_1, a_2, p_1, q_1, d_2, e_1, e_2 \in \mathbb{C}(a_2 \neq 0 \wedge F_1 = 0 \wedge F_2 = 0 \wedge F_3 = 0 \wedge F_3 = 0 \wedge F_4 = 0 \wedge F_5 = 0 \wedge$
$a_1^2 + a_2^2 + p_1 - 2a_1p_1 = 0 \wedge a_1^2 + a_2^2 + q_1 - 2a_1q_1 = 0 \Rightarrow P = 0).$

As a result, the condition $a_1^2 + a_2^2 - p_1 = 0 \wedge -2a_1 + a_1^2 + a_2^2 + q_1 = 0$ is a necessary and sufficient assumption for the conclusion. Note that the assumptions given in the original problem imply this condition, however, it is too severe. In order to get a natural interpretation of the condition we need a little bit of heuristics and light computations of real QE.

If we look at the problem carefully, we notice that P is on the right side of B and Q is left side of C, i.e. $p_1 \geq 0$ and $1 \geq q_1$. This is actually the desired assumption. In order to see it we need real QE computations. We can easily check the following sentences are true even by hand. (Actually we do not need the conditions $G_1 \geq 0$ and $G_2 \geq 0$.)

$\forall a_1, a_2, p_1 \in \mathbb{R}(a_2 \neq 0 \wedge G_1 \geq 0 \wedge a_2^2 + a_1^2 - p_1 = 0 \Rightarrow p_1 \geq 0).$

$\forall a_1, a_2, q_1 \in \mathbb{R}(a_2 \neq 0 \wedge G_2 \geq 0 \wedge -2a_1 + a_1^2 + a_2^2 + q_1 = 0 \Rightarrow 1 \geq q_1).$

We can also easily check the following formulas are equivalent to $a_1 = 1$ and $a_1 = 0$ respectively by any real QE program.

$\exists a_2, p_1 \in \mathbb{R}(a_2 \neq 0 \wedge G_1 \geq 0 \wedge a_1^2 + a_2^2 + p_1 - 2a_1p_1 = 0 \wedge p_1 \geq 0).$

$\exists a_2, q_1 \in \mathbb{R}(a_2 \neq 0 \wedge G_2 \geq 0 \wedge a_1^2 + a_2^2 + q_1 - 2a_1q_1 = 0 \wedge 1 \geq q_1 \geq 0).$

Note that $a_2^2 + a_1^2 - p_1$ and $a_1^2 + a_2^2 + p_1 - 2a_1p_1$ are identical when $a_1 = 1$, $-2a_1 + a_1^2 + a_2^2 + q_1$ and $a_1^2 + a_2^2 + q_1 - 2a_1q_1$ are also identical when $a_1 = 0$. Hence, the second sentence in the previous page is true.

In order to get a desired factorization, i.e., any factor is a linear polynomial of a target variable (p_1 and q_1 in the above example), we generally need some algebraic extension of $\mathbb{Q}$. We can also obtain such an algebraic extension by computation of Gröbner bases in general, although factorization of a polynomial ring over $\mathbb{Q}$ suffices for all problems of Hilbert geometry given in the past International Mathematical Olympiad.

5 Conclusion and Remarks

Many works on geometry theorem proving such as [2] have been done to date. Their main purpose is to develop advanced algorithms to get automatic proofs of difficult theorems. For example, the paper [10] gives an absolute factorization algorithm which efficiently computes the minimal algebraic extension. Meanwhile, we do not need such a sophisticated algorithm for our purpose as is mentioned in the last section. With a minimum knowledge given in this paper, anyone can apply our method in any CAS which can compute Gröbner bases, complex QE and real QE. Real QE can deal with any problem, however, its computation is very heavy and many problems cannot be handled by any of the existing real QE implementation. For a problem of metric geometry, complex QE is sufficient. Its computation is much faster than real QE. For a problem of Hilbert geometry, Gröbner basis computation, the fastest among others, is also useful for reducing an original problem to a much simpler subproblem which can be handled by most existing real QE programs. According to our computation experiment, we can detect unnecessary assumptions within a few seconds for any problem given in the past International Mathematical Olympiad using our method on Mathematica, Maple and REDUCE.

References

[1] The RegularChains Library. http://www.regularchains.org/index.html

[2] Chou, SC. Mechanical Geometry Theorem Proving. Mathematics and Its Applications, Springer, 1988.

[3] Collins, G, E. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. Automata theory and formal languages (Second GI Conf., Kaiserslautern, 1975), pp. 134-183, Lecture Notes in Comput. Sci., Vol. 33, Springer, Berlin, 1975.

[4] Fukasaku, R.,Inoue, S. and Sato, Y. On QE Algorithms over Algebraically Closed Field based on Comprehensive Gröbner Systems. Mathematics in Computer Science, Vol.9/3, 267-281, Birkhäuser/Springer, 2015.

[5] Fukasaku, R., Iwane, H. and Sato, Y. Real Quantifier Elimination by Computation of Comprehensive Gröbner Systems. Proceedings of ISSAC 2015, pp. 173-180, ACM, 2015.

[6] Mathematica Tutorial: tutorial/{ComplexPolynomialSystems,RealPolynomialSystems}

[7] QEPCAD-Quantifier Elimination by Partial Cylindrical Algebraic Decomposition. http://www.usna.edu/CS/qepcadweb/B/QEPCAD.html

[8] Redlog: an integral part of the interactive computer algebra system Reduce. http://www.redlog.eu/

[9] Tarski, A. A Decision Method for Elementary Algebra and Geometry. Quantifier Elimination and Cylindrical Algebraic Decomposition, Texts & Monographs in Symbolic Computation, pp. 24-84, Springer, 1998.

[10] Wang, D. Algebraic factoring and geometry theorem proving. In Automated Deduction — CADE-12 (1994) LNAI814 Springer,1994,386-400