Automated Geometric Theorem Proving

Presented by: Manish Patel

Institute: The Institute Of Science

Course: MSc Mathematics

File: mainbeamer

XeLaTeXed on 2022/12/21

Table of contents

Background Information

Different Geometries

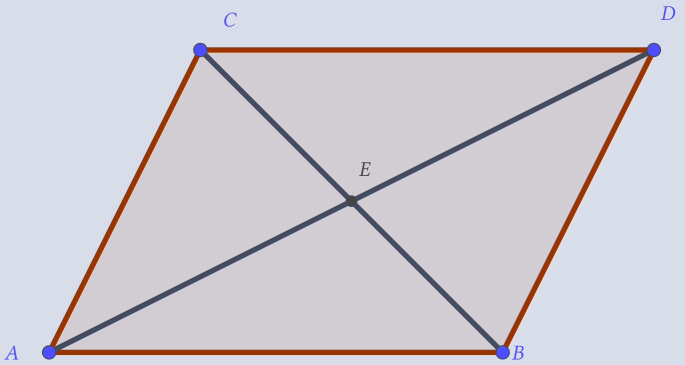
- Affine Geometry
- Metric Geometry
- Hilbert Geometry
- Tarski Geometry

Defferent Proving Strategies

Traditional Proof

Diagonals of a parallelogram bisect each other.

Let A, B, C, D be the vertices of a parallelogram in the plane. The two diagonals $\overline{AD}$ and $\overline{BC}$ of any parallelogram intersect at a point which bisects both diagonals.



Solution:

$\triangle ADB \cong \triangle DAC$

hence $AB \equiv CD$

$\triangle AEB \cong \triangle DEC$

hence $AE \equiv DE$

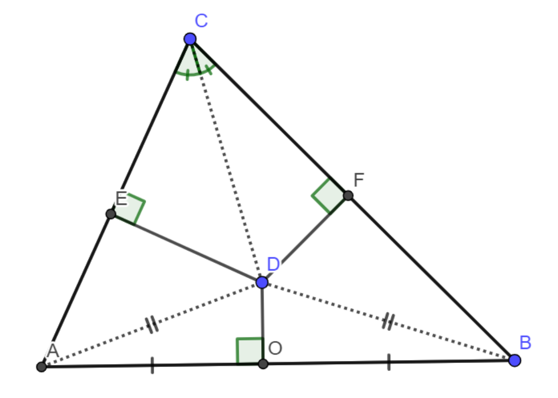
Therefore, diagonals of a parallelogram bisect each other.

We can let $A = (0,0), B = (u_1, 0), C = (u_2, u_3), D = (x_1, x_2)$, and $E = (x_3, x_4)$. Then we want to prove $g = x_1^2 - 2x_1x_3 - 2x_4x_2 + x_2^2$.

$$\begin{aligned} h_1 &= x_2 - u_3 \\ h_2 &= (x_1 - u_1) u_3 - u_2 x_2 \\ h_3 &= x_4 x_1 - x_3 x_2 \\ h_4 &= x_4 (u_2 - u_1) - (x_3 - u_1) u_3. \end{aligned}$$

$$g = x_1^2 - 2x_1x_3 - 2x_4x_2 + x_2^2.$$

Every triangle is isosceles. Let ABC be a triangle as shown in figure. We want to prove $CA \equiv CB$.



Proof. It is easy to see that $\triangle CDE \cong \triangle CDF$ and $\triangle ADE \cong \triangle BDF$. Hence $CE + EA = CF + FB$, i.e., $CA \equiv CB$

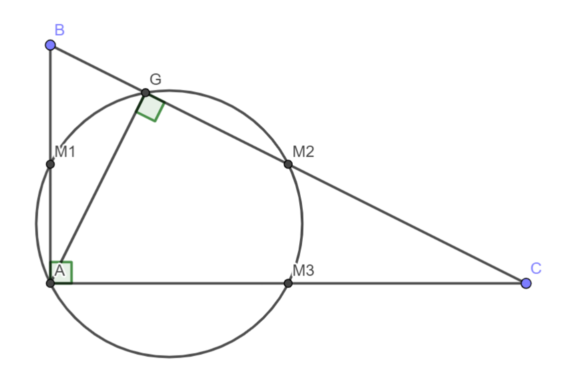
Geometric Config to Polynomials

Let A, B, C, D, E, F be points in the plane. Each of the following geometric statements can be expressed by one or more polynomial equations:

- $\overline{AB}$ is perpendicular to $\overline{CD}$.
- A, B, C are collinear.
- The distance from A to B is equal to the distance from C to D i.e. $AB = CD$.
- C lies on the circle with center A and radius AB.
- C is the midpoint of $\overline{AB}$.
- The acute angle $\angle ABC$ is equal to the acute angle $\angle DEF$
- $\overline{BD}$ bisects the angle $\angle ABC$.

Circle Theorem of Appolonius

Let $\triangle ABC$ be a right triangle in the plane, with right angle at A. The midpoints of the three sides and the foot of the altitude drawn from A to $\overline{BC}$ all lie on one circle.



We begin by constructing the triangle.

A at $(0,0)$ B at $(u_1, 0)$, the hypothesis that $\angle CAB$ is a right angle says $C = (0, u_2)$.

$M_1 = (x_1, 0)\,, M_2 = (0, x_2)$, and $M_3 = (x_3, x_4)$.

We obtain the equations

$$\begin{aligned}
h_1 &= 2x_1 - u_1 = 0,\\
h_2 &= 2x_2 - u_2 = 0,\\
h_3 &= 2x_3 - u_1 = 0,\\
h_4 &= 2x_4 - u_2 = 0.
\end{aligned}$$

The next step is to construct the point $H = (x_5, x_6)$, the foot of the altitude drawn from A. We have two hypotheses here:

$$\begin{aligned}
B, H, C \text{ collinear} : h_5 &= u_2 x_5 + u_1 x_6 - u_1 u_2 = 0,\\
AH \perp BC \,:\, h_6 &= u_1 x_5 - u_2 x_6 = 0.
\end{aligned}$$

Finally, we must consider the statement that M_1, M_2, M_3, H lie on a circle.

We call the center $O = (x_7, x_8)$ and derive two additional hypotheses:

$$M_1O = M_2O \,:\, h_7 = (x_1 - x_7)^2 + x_8^2 - x_7^2 - (x_8 - x_2)^2 = 0$$
$$M_1O = M_3O \,:\, h_8 = (x_1 - x_7)^2 + (0 - x_8)^2 - (x_3 - x_7)^2 - (x_4 - x_8)^2 = 0.$$

Our conclusion is $HO = M_1O$, which takes the form

$$g = (x_5 - x_7)^2 + (x_6 - x_8)^2 - (x_1 - x_7)^2 - x_8^2 = 0.$$

Generalization

For given geometric configuration, we will have some number of arbitrary coordinates, or independent variables in our construction, denoted by $u_1, \dots, u_m$. In addition, there will be some collection of dependent variables $x_1, \dots, x_n$.

The hypotheses of the theorem will be represented by a collection of polynomial equations in the u_i, x_j.

$$\begin{aligned} h_1\left(u_1, \dots, u_m, x_1, \dots, x_n\right) &= 0 \\ &\vdots \\ h_n\left(u_1, \dots, u_m, x_1, \dots, x_n\right) &= 0. \end{aligned}$$

The conclusions of the theorem will also be expressed as polynomials in the u_i, x_j.

$$g\left(u_1, \dots, u_m, x_1, \dots, x_n\right) = 0$$

How can the fact that g follows from $h_1, \dots, h_n$ be deduced algebraically?

The basic idea is that we want g to vanish whenever $h_1, \dots, h_n$ do.

Follows strictly

The conclusion g follows strictly from the hypotheses $h_1, \dots, h_n$ if $g \in \mathbf{I}(V) \subseteq \mathbb{R}[u_1, \dots, u_m, x_1, \dots, x_n]$, where $V = \mathbf{V}(h_1, \dots, h_n) \subseteq \mathbb{R}^{m+n}$

If $g \in \sqrt{\langle h_1, \dots, h_n \rangle} \subseteq \mathbf{I}(V)$, then g follows strictly from $h_1, \dots, h_n$.
Note that the converse fails whenever $\sqrt{\langle h_1, \dots, h_n \rangle} \subsetneq \mathbf{I}(V)$

Let $\bar{I} = \langle h_1, \dots, h_n, 1 - yg \rangle$ in the ring $\mathbb{R}[u_1, \dots, u_m, x_1, \dots, x_n, y]$, then

$$g \in \sqrt{\langle h_1, \dots, h_n \rangle} \iff \{1\} \text{ is the reduced Gröbner basis of } \bar{I}.$$

How can the fact that g follows from $h_1, \ldots, h_n$ be deduced algebraically?

$$g \in \sqrt{\langle h_1, \ldots, h_n \rangle} \subseteq \mathbb{R}[u_1, \ldots, u_m, x_1, \ldots, x_n] \iff g \in \mathbf{I}(V_{\mathbb{C}}) = \sqrt{\langle h_1, \ldots, h_n \rangle} \subseteq \mathbb{C}[u_1, \ldots, u_m, x_1, \ldots, x_n]$$

Example 1 (continued). Taking as hypotheses the four polynomials:

$$\begin{aligned}
h_1 &= x_2 - u_3 \\
h_2 &= (x_1 - u_1) u_3 - u_2 x_2 \\
h_3 &= x_4 x_1 - x_3 x_2 \\
h_4 &= x_4 (u_2 - u_1) - (x_3 - u_1) u_3.
\end{aligned}$$

We will take as conclusion the first polynomial:

$$g = x_1^2 - 2x_1 x_3 - 2x_4 x_2 + x_2^2.$$

How can the fact that g follows from $h_1, \dots, h_n$ be deduced algebraically?

Now must compute a Gröbner basis for

$$\bar{I} = \langle h_1, h_2, h_3, h_4, 1 - yg \rangle \subseteq \mathbb{R}\left[u_1, u_2, u_3, x_1, x_2, x_3, x_4, y\right].$$

Surprisingly enough, we do not find $\{1\}$.

Gröbner basis for $I = \langle h_1, h_2, h_3, h_4 \rangle$ in $\mathbb{R}[u_1, u_2, u_3, x_1, x_2, x_3, x_4]$, using lex order with $x_1 > x_2 > x_3 > x_4 > u_1 > u_2 > u_3$. The result is

$$
\begin{aligned}
f_1 &= x_1 x_4 + x_4 u_1 - x_4 u_2 - u_1 u_3, \\
f_2 &= x_1 u_3 - u_1 u_3 - u_2 u_3, \\
f_3 &= x_2 - u_3, \\
f_4 &= x_3 u_3 + x_4 u_1 + x_4 u_2 - u_1 u_3, \\
f_5 &= x_4 u_1^2 - x_4 u_1 u_2 - \frac{1}{2} u_1^2 u_3 + \frac{1}{2} u_1 u_2 u_3, \\
f_6 &= x_4 u_1 u_3 - \frac{1}{2} u_1 u_3^2.
\end{aligned}
$$

$V = \mathbf{V}(h_1, h_2, h_3, h_4) = \mathbf{V}(f_1, \ldots, f_6)$ in $\mathbb{R}^7$

Note f_2 factors as $(x_1 - u_1 - u_2) u_3$

$$V = \mathbf{V}(f_1, x_1 - u_1 - u_2, f_3, f_4, f_5, f_6) \cup \mathbf{V}(f_1, u_3, f_3, f_4, f_5, f_6)$$

Since f_5 and f_6 also factor, we can continue this decomposition process.

$$V = V' \cup U_1 \cup U_2 \cup U_3$$

into irreducible varieties, where

$$\begin{aligned}
V' &= \mathbf{V}\left(x_1 - u_1 - u_2, x_2 - u_3, x_3 - \frac{u_1 + u_2}{2}, x_4 - \frac{u_3}{2}\right), \\
U_1 &= \mathbf{V}(x_2, x_4, u_3) \\
U_2 &= \mathbf{V}(x_1, x_2, u_1 - u_2, u_3) \\
U_3 &= \mathbf{V}(x_1 - u_2, x_2 - u_3, x_3 u_3 - x_4 u_2, u_1)
\end{aligned}$$

$(u_3 = 0, x_2 = 0, x_4 = 0)$ here h_i's are simultaneously zero but g is not

Let $V = \mathbf{V}(h_1, \dots, h_n) \subseteq \mathbb{R}^{m+n}$ as a finite union of irreducible varieties

$$V = V_1 \cup \cdots \cup V_k.$$

Definition

Let W be an irreducible variety in the affine space $\mathbb{R}^{m+n}$ with coordinates $u_1, \dots, u_m, x_1, \dots, x_n$. We say that the functions $u_1, \dots, u_m$ are algebraically independent on W if $\mathbf{I}(W) \cap \mathbb{R}[u_1, \dots, u_m] = \{0\}$.

We can regroup the irreducible components in the following way:

$$V = W_1 \cup \cdots \cup W_p \cup U_1 \cup \cdots \cup U_q,$$

$$V' = W_1 \cup \cdots \cup W_p \subseteq V.$$

Follows Generically

Definition

The conclusion g follows generically from the hypotheses $h_1, \ldots, h_n$ if $g \in \mathbf{I}(V') \subseteq \mathbb{R}[u_1, \ldots, u_m, x_1, \ldots, x_n]$, where, as above, $V' \subseteq \mathbb{R}^{m+n}$ is the union of the components of the variety $V = \mathbf{V}(h_1, \ldots, h_n)$ on which the u_i are algebraically independent.

The conclusion g follows generically from $h_1, \ldots, h_n$ whenever there is some nonzero polynomial $c(u_1, \ldots, u_m) \in \mathbb{R}[u_1, \ldots, u_m]$ such that

$$c \cdot g \in \sqrt{H},$$

where H is the ideal generated by the hypotheses h_i in $\mathbb{R}[u_1, \ldots, u_m, x_1, \ldots, x_n]$.

Follows Generically

The following are equivalent:

1. There is a nonzero polynomial $c \in \mathbb{R}[u_1, \dots, u_m]$ such that $c \cdot g \in \sqrt{H}$.
2. $g \in \sqrt{\widetilde{H}}$, where $\widetilde{H}$ is the ideal generated by the h_j in $\mathbb{R}(u_1, \dots, u_m)[x_1, \dots, x_n]$.
3. $\{1\}$ is the reduced Gröbner basis of the ideal

$$\langle h_1, \dots, h_n, 1 - yg \rangle \subseteq \mathbb{R}(u_1, \dots, u_m)[x_1, \dots, x_n, y]$$

We will call this the Gröbner basis method in geometric theorem proving.

We first compute a Gröbner basis of the ideal $\langle h_1, h_2, h_3, h_4, 1 - yg \rangle$ in the ring $\mathbb{R}(u_1, u_2, u_3)[x_1, x_2, x_3, x_4, y]$. This computation does yield $\{1\}$ as we expect.

Pseudodivision

Let $f, g \in k[x_1, \dots, x_n, y]$ and assume $m \leq p$ and $d_m \neq 0$.

$$f = c_p y^p + \cdots + c_1 y + c_0,$$
$$g = d_m y^m + \cdots + d_1 y + d_0,$$

There is an equation

$$d_m^s f = qg + r,$$

where $q, r \in k[x_1, \dots, x_n, y]$, $s \geq 0$, and either $r = 0$ or the degree of r in y is less than m.

$r \in \langle f, g \rangle$ in the ring $k[x_1, \dots, x_n, y]$.

For example, if we pseudodivide $f = x^2y^3 - y$ by $g = x^3y - 2$ with respect to y by the algorithm above, we obtain the equation

$$\left(x^3\right)^3 f = \left(x^8y^2 + 2x^5y + 4x^2 - x^6\right) g + 8x^2 - 2x^6.$$

In particular, the pseudoremainder is $\mathrm{Rem}(f, g, y) = 8x^2 - 2x^6$.

Wu's Method[1]

Step 1. Conversion of a geometry statement into the corresponding polynomial equations.

Step 2. Triangulation of the hypothesis polynomials using pseudo division.

$$
\begin{aligned}
f_1 &= f_1\left(u_1, \dots, u_m, x_1\right) \\
f_2 &= f_2\left(u_1, \dots, u_m, x_1, x_2\right) \\
&\vdots \\
f_n &= f_n\left(u_1, \dots, u_m, x_1, \dots, x_n\right)
\end{aligned}
$$

[1] https://mybinder.org/v2/gh/primepatel/Bhumiti/HEAD?labpath=demo.ipynb

Step 3. Successive pseudo division to compute the final remainder R_0.

$$\begin{aligned}
R_{n-1} &= \operatorname{Rem}(g, f_n, x_n)\,, \\
R_{n-2} &= \operatorname{Rem}(R_{n-1}, f_{n-1}, x_{n-1})\,, \\
&\vdots \\
R_1 &= \operatorname{Rem}(R_2, f_2, x_2)\,, \\
R_0 &= \operatorname{Rem}(R_1, f_1, x_1)\,.
\end{aligned}$$

Step 4. Analysis of nondegenerate conditions $d_1 \neq 0, \ldots, d_r \neq 0$

Main Idea

If $R_0 = 0$ in $d_1^{s_1} \cdots d_n^{s_n} g = A_1 f_1 + \cdots + A_n f_n + R_0$

$$h_1 = 0 \wedge \ldots \wedge h_n = 0 \wedge d_1 \neq 0 \wedge \ldots \wedge d_k \neq 0 \Rightarrow g = 0$$

Affine Space

Affine Space

Given a field k and a positive integer n, we define the n-dimensional affine space over k to be the set

$$k^n = \{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in k\}$$

Affine Varieties

Let k be a field, and let $f_1, \ldots, f_s$ be polynomials in $k[x_1, \ldots, x_n]$. Then we set

$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in k^n \mid f_i(a_1, \ldots, a_n) = 0 \text{ for all } 1 \leq i \leq s\}.$$

We call $\mathbf{V}(f_1, \ldots, f_s)$ the affine variety defined by $f_1, \ldots, f_s$.

Affine Space

In the plane $\mathbb{R}^2$ with the variety $\mathbf{V}\left(x^2+y^2-1\right)$, which is the circle of radius 1 centered at the origin

If $V, W \subseteq k^n$ are affine varieties, then so are $V \cup W$ and $V \cap W$.

Suppose that $V = \mathbf{V}(f_1, \ldots, f_s)$ and $W = \mathbf{V}(g_1, \ldots, g_t)$.

$$V \cap W = \mathbf{V}(f_1, \ldots, f_s, g_1, \ldots, g_t)$$
$$V \cup W = \mathbf{V}\left(f_i g_j \mid 1 \le i \le s, 1 \le j \le t\right)$$

$\mathbf{I}(V) = \{f \in k[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0 \text{ for all } (a_1, \ldots, a_n) \in V\}$ is ideal

$$\overset{polynomials}{f_1, \cdots f_s} \to \overset{variety}{V(f_1, \cdots f_s)} \to \overset{Ideal}{I(V(f_1, \cdots f_s))}$$

Affine Space

Hilbert's Nullstellensatz

Let k be an algebraically closed field. If $f, f_1, \dots, f_s \in k[x_1, \dots, x_n]$, then $f \in \mathbf{I}(\mathbf{V}(f_1, \dots, f_s))$ if and only if

$$f^m \in \langle f_1, \dots, f_s \rangle$$

for some integer $m \geq 1$.

Radical Membership

Let k be an arbitrary field and let $I = \langle f_1, \dots, f_s \rangle \subseteq k[x_1, \dots, x_n]$ be an ideal. Then $f \in \sqrt{I}$ if and only if the constant polynomial 1 belongs to the ideal $\tilde{I} = \langle f_1, \dots, f_s, 1 - yf \rangle \subseteq k[x_1, \dots, x_n, y]$, in which case $\tilde{I} = k[x_1, \dots, x_n, y]$.

back to there

References

References

Dr. Wen-tsün Wu
Mechanical Theorem Proving in Geometries: Basic Principles
Springer, 1994

Shang-Ching Chou
Mechanical Geometry Theorem Proving
Springer, 1987

David A. Cox, John Little, Donal O'Shea
Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra
Springer, 2007

Dr. Dongming Wang
Elimination Methods
Springer, 2001