Gröbner Bases Tutorial

Part II: A Sampler of Recent Developments

David A. Cox

Department of Mathematics and Computer Science
Amherst College
`dac@cs.amherst.edu`

ISSAC 2007 Tutorial

Outline

The original plan was to cover five topics:

- Graph Theory
- Geometric Theorem Discovery
- The Generic Gröbner Walk
- Alternatives to the Buchberger Algorithm ← *Too hard*
- Moduli of Quiver Representations ← *Too complicated*

The new plan is to cover four topics:

- Graph Theory
- Geometric Theorem Discovery
- The Generic Gröbner Walk
- (New) Phylogenetic Invariants

Graph Colorings

Let $G = (V, E)$ be a graph with vertices $V = \{1, \dots, n\}$.

Definition

A *k*-coloring of G is a function from V to a set of k colors such that adjacent vertices have distinct colors.

Example

vertices $=$ 81 squares
edges $=$ links between:

- squares in same column
- squares in same row
- squares in same 3×3

Colors $= \{1, 2, \dots, 9\}$

Goal: Extend the partial coloring to a full coloring.

				3	5			
	1		2			9		
7		6				2		
6			5				3	
2				4				9
	3				1			5
		3				4		8
		4			6		7	
			3	1				

Graph Ideal

Definition

The *k*-coloring ideal of G is the ideal $I_{G,k} \subseteq \mathbb{C}[x_i \mid i \in V]$ generated by:

$$\begin{aligned} \text{for all } i \in V: &\quad x_i^k - 1 \\ \text{for all } ij \in E: &\quad x_i^{k-1} + x_i^{k-2}x_j + \cdots + x_i x_j^{k-2} + x_j^{k-1}. \end{aligned}$$

Lemma

$\mathbf{V}(I_{G,k}) \subseteq \mathbb{C}^n$ *consists of all* k*-colorings of* G *for the set of colors consisting of the* k^{th} *roots of unity*

$$\mu_n = \{1,\ \zeta_k,\ \zeta_k^2,\ \ldots\ ,\ \zeta_k^{k-1}\}, \quad \zeta_k = e^{2\pi i/k}.$$

Proof. $$\frac{(x_i^k - 1) - (x_j^k - 1)}{x_i - x_j} = x_i^{k-1} + x_i^{k-2}x_j + \cdots + x_j^{k-1}.$$ □

The Existence of Colorings

Two Observations

- G has a k-coloring $\iff \mathbf{V}(I_{G,k}) \neq \emptyset$.
- Hence the Consistency Theorem gives a Gröbner basis criterion for the existence of a k-coloring.

3-Colorings

For 3-colorings, the ideal $I_{G,3}$ is generated by

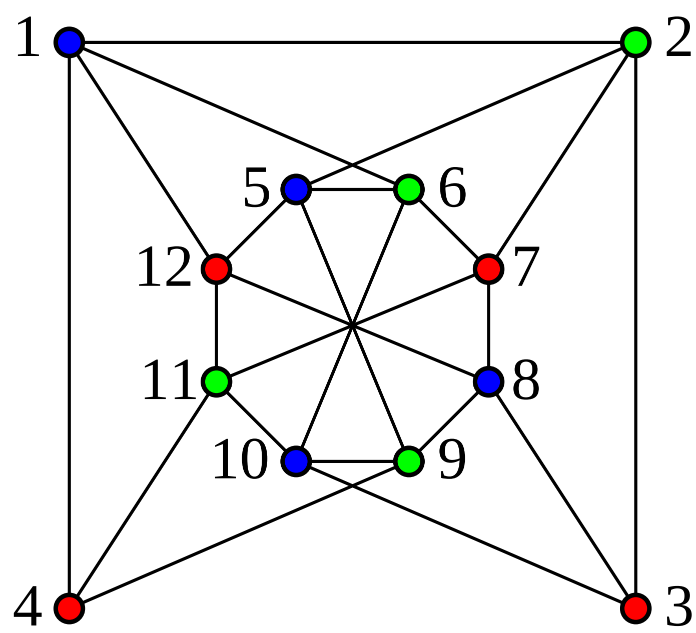
$$\begin{aligned} \text{for all } i \in V: &\quad x_i^3 - 1 \\ \text{for all } ij \in E: &\quad x_i^2 + x_i x_j + x_j^2. \end{aligned}$$

These equations can be hard to solve!

Theorem

3-colorability is NP-complete.

Example



This example of a graph with a 3-coloring is due to Chao and Chen (1993).

Hillar and Windfeldt (2006) compute the reduced Gröbner basis of the graph ideal $I_{G,3}$ for lex with $x_1 > \cdots > x_{12}$.

The reduced Gröbner basis is:

$$\{x_{12}^3 - 1,\ x_7 - x_{12},\ x_4 - x_{12},\ x_3 - x_{12},\ x_{11}^2 + x_{11}x_{12} + x_{12}^2,\ x_9 - x_{11},\ x_6 - x_{11},\ x_2 - x_{11},\ x_{10} + x_{11} + x_{12},\ x_8 + x_{11} + x_{12},\ x_5 + x_{11} + x_{12},\ x_1 + x_{11} + x_{12}\}.$$

Note $x_8 - x_{10},\ x_5 - x_{10},\ x_1 - x_{10} \in I_{G,3}$.

Uniquely k-Colorable Graphs

The Chao/Chen graph has essentially only one 3-coloring.

Definition

A graph G is uniquely k-colorable if it has a unique k-coloring up the permutation of the colors.

Hillar and Windfeldt show that unique k-colorability is easy to detect using Gröbner bases.

We start with a k-coloring of G that uses all k colors. Assume the k colors occur among the last k vertices. Then:

- Use variables $x_1, \ldots, x_{n-k}, y_1, \ldots, y_k$ with lex order

$$x_1 > \cdots > x_{n-k} > y_1 > \cdots > y_k.$$

- Use these variables to label the vertices of G.

Some Interesting Polynomials

Consider the following polynomials:

$$y_k^k - 1$$

$$h_j(y_j,\dots,y_k) = \sum_{\alpha_j+\cdots+\alpha_k=j} y_j^{\alpha_j}\cdots y_k^{\alpha_k}, \quad j=1,\dots,k-1$$

$$x_i - y_j, \quad \text{color}(x_i) = \text{color}(y_j),\ j \geq 2$$

$$x_i + y_2 + \cdots + y_k, \quad \text{color}(x_i) = \text{color}(y_1).$$

In this notation, the Gröbner basis given earlier is:

$$\{y_3^3 - 1,$$
$$h_2(y_2,y_3) = y_2^2 + y_2y_3 + y_3^2,\ h_1(y_1,y_2,y_3) = y_1 + y_2 + y_3,$$
$$x_7 - y_3,\ x_4 - y_3,\ x_3 - y_3,\ x_9 - y_2,\ x_6 - y_2,\ x_2 - y_2,$$
$$x_8 + y_2 + y_3,\ x_5 + y_2 + y_3, x_1 + y_2 + y_3\}.$$

A Theorem

Summary:

- G has vertices $x_1,\ldots,x_{n-k}, y_1,\ldots,y_k$.
- G has a k-coloring where $y_1,\ldots,y_k$ get all the colors.
- $\mathbb{C}[\mathbf{x},\mathbf{y}]$ has lex with $x_1 > \cdots > x_{n-k} > y_1 > \cdots > y_k$.

Using this data, we create:

- The coloring ideal $I_{G,k} \subseteq \mathbb{C}[\mathbf{x},\mathbf{y}]$.
- The n polynomials $g_1,\ldots,g_n$ given by
  $y_k^k - 1,\ \ h_j(y_j,\ldots,y_k),\ \ x_i - y_j,\ \ x_i + y_2 + \cdots + y_k,$

Theorem

The following are equivalent:

- *G is uniquely k-colorable.*
- $g_1,\ldots,g_n \in I_{G,k}$.
- $\{g_1,\ldots,g_n\}$ *is the reduced Gröbner basis for* $I_{G,k}$.

Remarks

When G is uniquely k-colorable, the theorem implies

$$\langle \mathrm{LT}(I_{G,k})\rangle = \langle y_k^k,\, y_{k-1}^{k-1},\, \ldots\, ,\, y_2^2,\, y_1,\, x_{n-k},\, \ldots\, ,\, x_1\rangle.$$

Since

- $\dim \mathbb{C}[\mathbf{x},\mathbf{y}]/I_{G,k} = \#$monomials not in $\langle \mathrm{LT}(I_{G,k})\rangle$, and
- $I_{G,k}$ is radical,

a uniquely k-colorable graph has

$$\#k\text{-colorings} = \dim \mathbb{C}[\mathbf{x},\mathbf{y}]/I_{G,k} = k\cdot(k-1)\cdots 2\cdot 1 = k!.$$

Hillar and Windfeldt have a version of this result that doesn't assume we know a k-coloring in advance.

Final Comment

Other aspects of graphs can be coded algebraically. Here is an example from de Loera, Lee, Margulies, and Onn (2007).

Let $G = (V, E)$ with $V = \{1, \dots, n\}$. Consider variables $x_1, \dots, x_n$ and $y_1, \dots, y_n$ and fix a positive integer L.

Theorem

G has a cycle of length L $\iff$ the following equations have a solution:

$$\begin{aligned} y_1 + \cdots + y_n &= L \\ y_i(y_i - 1) &= 0, \quad 1 \le i \le n \\ \textstyle\prod_{s=1}^{n}(x_i - s) &= 0, \quad 1 \le i \le n \\ y_i \prod_{ij \in E} (x_i - y_j x_i + y_j)(x_i - y_j x_i - y_j(L-1)) &= 0, \quad 1 \le i \le n. \end{aligned}$$

A Final Amusement

To solve this sudoku, use:

- 81 variables x_{ij}, $1 \leq i,j \leq 9$.
- Relabel the 9 variables for red squares as $y_1,\ldots,y_9$.
- The graph ideal $I_{G,9}$.
- The 9 polynomials $y_9{}^9 - 1$, $h_8(y_8,y_9), h_7(y_7,y_8,y_9)$, $h_6(y_6,y_7,y_8,y_9),\ldots$, $h_1(y_1,\ldots,y_9) = y_1 + \cdots + y_9$.
- The 16 polynomials $x_{31} - y_7$, $x_{33} - y_6, x_{37} - y_2, \ldots$

				3	5			
	1		2			9		
7		6				2		
6			5				3	
2				4				9
	3				1			5
		3				4		8
		4			6		7	
			3	1				

Assuming a unique solution, the Gröbner basis of the ideal generated by these polynomials will contain $x_{11} - y_i$, etc. This will tell us how to fill in the blank squares!

References

- C.-Y. Chao, Z. Chen, *On uniquely 3-colorable graphs*, Discrete Mathematics **112** (1993), 21–27.
- C. Hillar, T. Windfeldt, *Algebraic characterization of uniquely vertex colorable graphs*, J. Comb. Th., to appear.
- J. A. de Loera, *Gröbner bases and graph colorings*, Beiträge zur Algebra un Geometrie **36** (1995), 89–96.
- J. A. de Loera, J. Lee, S. Margulies, S. Onn, *Expressing combinatorial optimization problems by systems of polynomial equations and the Nullstellensatz*, arXiv:0706.0578 [math.CO].

Geometric Theorem Discovery

Our next topic involves an application of comprehensive Gröbner systems to the problem of discovering the correct hypotheses that give an interesting theorem in geometry.

Our discussion was inspired by a preprint of Montes and Recio (2007).

We begin with an example of Sato and Suzuki (2006) that illustrates specialization of Gröbner bases.

Example 1

The ideal $I := \langle (u-1)x + y^2, uy + u \rangle \subseteq k[x,y,u]$. A lex Gröbner basis for $x > y > u$ is

$$\{ux - x + y^2, uy + u, xy + x - y^3 - y^2\}$$

We think of u as a parameter.

Set $V := \mathbf{V}(I) \subseteq \mathbb{A}^3$. Let's apply our theorems:

- Elimination Theorem $\Rightarrow I_2 = \{0\}$.
- Closure Theorem $\Rightarrow$ the projection of V onto the last coordinate has Zariski dense image in $\mathbb{A}^1$.

A more careful analysis reveals that $\pi_2(V) = \mathbb{A}^1 \setminus \{1\}$.

This raises two questions.

First Question

Does $G = \{ux - x + y^2,\ uy + u,\ xy + x - y^3 - y^2\}$ remain a Gröbner basis (for lex with $x > y$) when the parameter u is given a specific numerical value b?

Two observations:

- Setting $u = 1$ gives $\overline{G}_1 = \{y^2,\ y + 1,\ xy + x - y^3 - y^2\}$, which generates $\langle 1 \rangle = k[x, y]$. Since $1 \notin \langle y^2, y, xy \rangle$, $\overline{G}$ is not a Gröbner basis.
- Write G as

$$\{(u-1)\cdot x + y^2,\ u \cdot y + u,\ 1 \cdot xy + x - y^3 - y^2\}.$$

  If $u = b \neq 0, 1$, the Special Case considered in the proof of the Closure Theorem implies that $\overline{G}_b$ is a Gröbner basis.

General Question: How do Gröbner bases specialize?

Second Question

Since $\pi_2(V) = \mathbb{A}^1 \setminus \{1\}$, the equations

$$(u-1)x + y^2 = uy + u = 0$$

have a solution when $u = b \neq 1$.

How many solutions?

- $u = b \neq 0, 1 \Rightarrow (b-1)x + y^2 = y + 1 = 0$ has a unique solution.
- $u = 0 \Rightarrow -x + y^2 = 0$ has infinitely many solutions.

General Question: How do we describe the number of solutions?

Answers for the Example

Consider the following pairs:

$$(S_1, G_1) := \big(\mathbb{A} \setminus \{0,1\}, \{(u-1)x + y^2, uy + u\}\big)$$
$$(S_2, G_2) := \big(\{0\}, \{x - y^2\}\big)$$
$$(S_3, G_3) := \big(\{1\}, \{1\}\big).$$

Note that:

- S_i is constructible, $S_1 \cup S_2 \cup S_3 = \mathbb{A}^1$ is a partition.
- For $b \in S_i$, $\overline{G_i}_b$ is a reduced Gröbner basis (almost).
- For $b \in S_i$, $\langle \mathrm{LT}(\overline{G_i}_b) \rangle$ is independent of b.
- $\langle \mathrm{LT}(\overline{G_1}_b) \rangle = \langle x, y \rangle$, $\langle \mathrm{LT}(\overline{G_2}_b) \rangle = \langle x \rangle$, $\langle \mathrm{LT}(\overline{G_3}_b) \rangle = \langle 1 \rangle$ gives the number of solutions.

This is a minimal canonical comprehensive Gröbner system.

MCCGS

Let $I \subseteq k[\mathbf{x},\mathbf{u}]$ be an ideal with variables $\mathbf{x} = (x_1,\ldots,x_n)$ and parameters $\mathbf{u} = (u_1,\ldots,u_m)$. Fix an order $>$ on $k[\mathbf{x}]$.

Definition

A minimal canonical comprehensive Gröbner system for I and $>$ consists of pairs (S_i, G_i) satisfying:

- The S_i give a constructible partition of $\mathbb{A}^m$.
- For $\mathbf{b} \in S_i$, setting $\mathbf{u} = \mathbf{b}$ gives a reduced Gröbner basis $\overline{G_{i\mathbf{b}}}$ (up to constants).
- For $\mathbf{b} \in S_i$, $\langle \mathrm{LT}(\overline{G_{i\mathbf{b}}})\rangle$ is independent of $\mathbf{b}$.
- No smaller partition exists with these properties.

This definition is due to Manubens and Montes (2006).

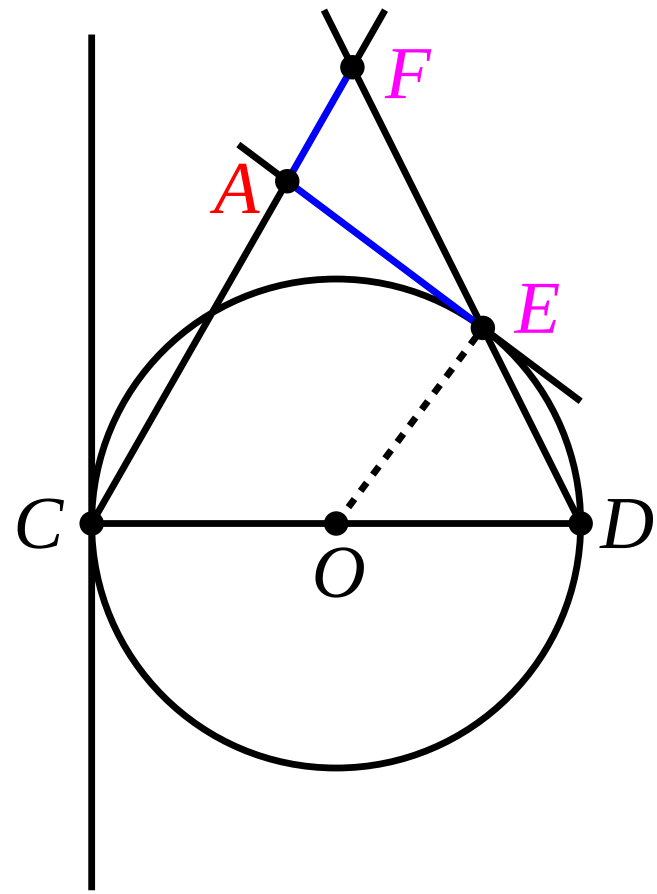
A False Theorem

Let CD be the diameter of a circle of radius 1. Fix A. Then:

- The line $\overleftrightarrow{AE}$ is tangent to the circle at E.
- The lines $\overleftrightarrow{AC}$ and $\overleftrightarrow{ED}$ meet at F.

False Theorem

$AE = AF$.



Challenge: Discover reasonable hypotheses on A to make the theorem true.

Hypotheses

Set $A = (u_1, u_2)$
$E = (x_1, x_2)$
$F = (x_3, x_4)$.

Then:

- $\overleftrightarrow{AE} \perp \overleftrightarrow{OE}$ gives

  $h_1 := (x_1 - u_1)(x_1 - 1) + (x_2 - u_2)x_2$.

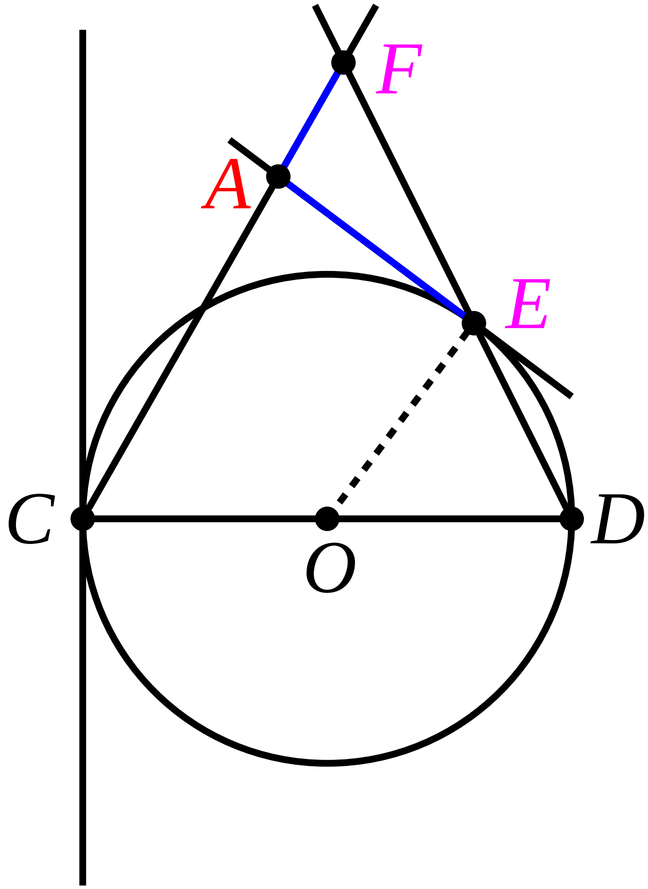
- $OE = 1$ gives

  $h_2 := (x_1 - 1)^2 + x_2^2 - 1$.

- $F = \overleftrightarrow{AC} \cap \overleftrightarrow{ED}$ gives

  $h_3 := u_1 x_4 - u_2 x_3$.

  $h_4 := x_4(x_1 - 2) - x_2(x_3 - 2)$.



More Hypotheses

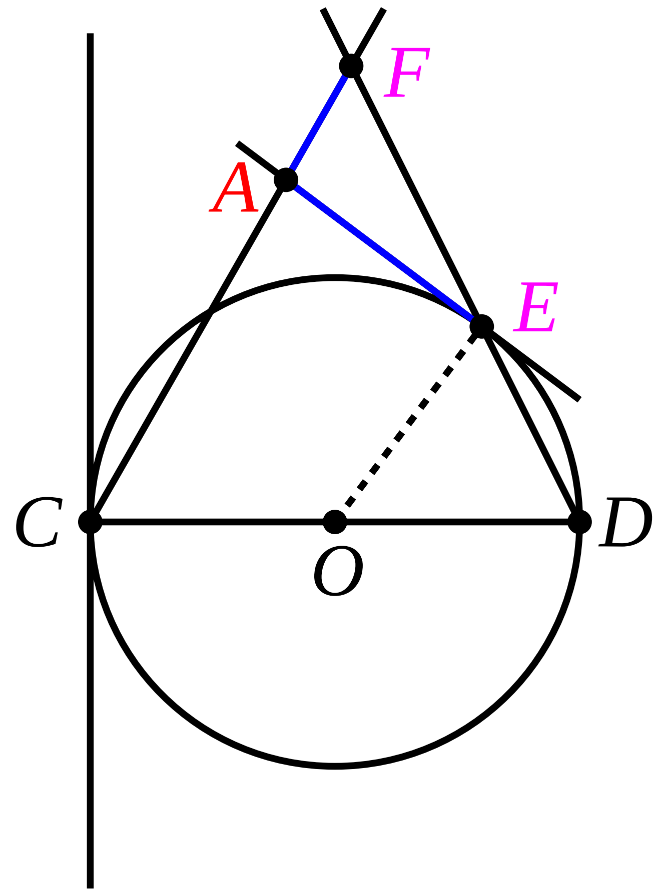
We also need to assume:

- $A \neq C$, so

  $u_1 \neq 0$ or $u_2 \neq 0$.

- $E \neq D$, so

  $x_2 \neq 2$.



Conclusion: The ideal that describes this problem is the saturation

$$I := \langle h_1, h_2, h_3, h_4 \rangle : \langle (x_2 - 2)u_1, (x_2 - 2)u_2 \rangle^\infty$$

in the ring $k[x_1, x_2, x_3, x_4, u_1, u_2]$.

Strategy

Our false theorem asserts $AE = AF$. This gives

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

Strategy

Compute a MCCGS for the ideal

$$I + \langle g \rangle \subseteq k[x_1, x_2, x_3, x_4, u_1, u_2], \quad u_1, u_2 \text{ parameters.}$$

Intuition

The false theorem is true for those $\mathbf{u} = \mathbf{b} \in \mathbb{A}^2$ for which

$$\emptyset \neq \mathbf{V}(\bar{I}_{\mathbf{b}} + \langle \bar{g}_{\mathbf{b}} \rangle) \subseteq \mathbb{A}^4.$$

The MCCGS

The MCCGS for $I + \langle g \rangle \subseteq k[x_1, x_2, x_3, x_4, u_1, u_2]$ under lex order with $x_1 > x_2 > x_3 > x_4$ is

$$(S_1, G_1) \cup \cdots \cup (S_6, G_6)$$

The S_i and Leading Terms

i	S_i	$\mathrm{LT}(\overline{G}_{i\,\mathbf{b}})$
1	$\mathbb{A}^2 \setminus \big(\mathbf{V}(u_1^2 + u_2^2 - 2u_1) \cup \mathbf{V}(u_1)\big)$	1
2	$\mathbf{V}(u_1^2 + u_2^2 - 2u_1) \setminus \{(0,0),(2,0)\}$	x_1, x_2, x_3, x_4^2
3	$\mathbf{V}(u_1) \setminus \{(0,0),(0,\pm i)\}$	x_1, x_2, x_3, x_4^2
4	$\{(0,\pm i)\}$	x_1, x_2, x_3, x_4
5	$\{(2,0)\}$	x_1, x_2^2, x_3, x_4^2
6	$\{(0,0)\}$	x_1, x_2, x_3^2, x_4^2

Consequences of the MCCGS

- We can ignore

$$S_4 = \{(0,\pm i)\},\ \ S_5 = \{(2,0)\},\ \ S_6 = \{(0,0)\}.$$

  The first is not real, and the second and third are impossible since $E \neq D$ and $A \neq C$.
- $G_1 = \{1\}$ on $S_1 = \mathbb{A}^2 \setminus \big(\mathbf{V}(u_1^2+u_2^2-2u_1) \cup \mathbf{V}(u_1)\big) \Rightarrow$

$$\mathbf{V}(I + \langle g \rangle) = \emptyset \ \text{ if } \ \mathbf{u} = \mathbf{b} \notin \mathbf{V}(u_1^2+u_2^2-2u_1) \cup \mathbf{V}(u_1).$$

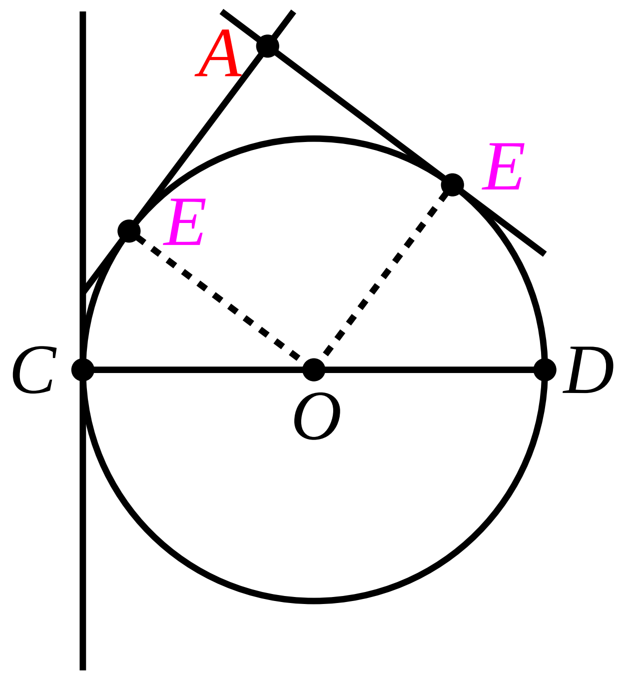
- Hence the "false theorem" $AE = AF$ (i.e., $g = 0$) cannot follow from our hypotheses (i.e., the ideal I) unless the point A comes from $\mathbf{V}(u_1^2+u_2^2-2u_1) \cup \mathbf{V}(u_1)$.
- This holds $\Leftrightarrow$ A is on the circle or the tangent at C.

Consequences and Expectations

Consequence

"A is on the circle or the tangent at C" is a necessary condition for $AE = AF$.

Before we investigate sufficiency, note that when a solution exists, we expect two solutions: Given A, there are two choices for E:



S_2 and S_3

To study whether the condition:

"A is on the circle or the tangent at C"

is sufficient, we use:

$$S_2 = \mathbf{V}(u_1^2 + u_2^2 - 2u_1) \setminus \{(0,0),(2,0)\}$$
$$S_3 = \mathbf{V}(u_1) \setminus \{(0,0),(0,\pm i)\}.$$

Both have $\mathrm{LT}(\overline{G_{i\,\mathbf{b}}}) = \langle x_1, x_2, x_3, x_4^2 \rangle$, so there are two solutions (counting multiplicity).

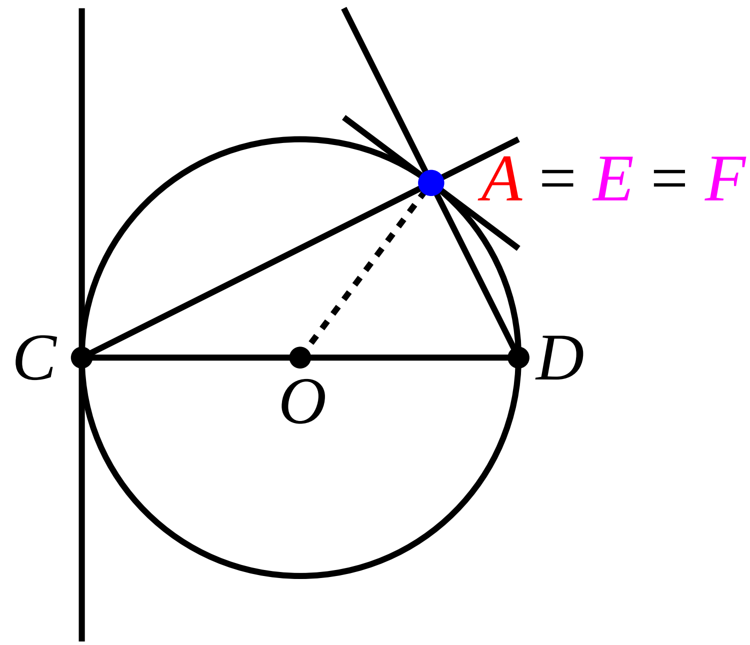
Notice that:

- S_2 corresponds to "A is on the circle".
- S_3 corresponds to "A is on the tangent at C".

We study each case separately.

On the Circle

When A is on the circle, we get:



Here, $AE = AF$ is true – both sides are zero! The unique solution has multiplicity two (the tangents from A coincide).

Hence the “false theorem” is true but not interesting.

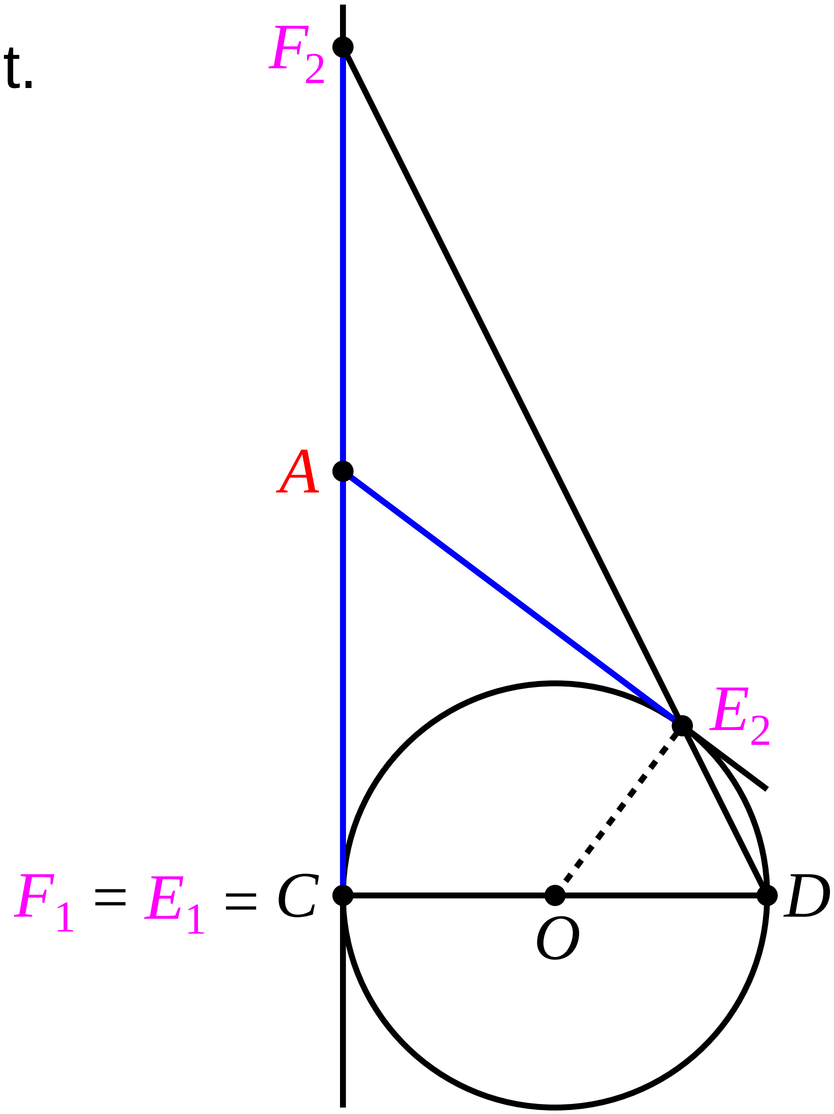
On the Tangent

When A is on the tangent, we get the picture to the right.

There are two choices for E:

- For E_1, we get $F_1 = E_1$, so $AE = AF$ is true but uninteresting.
- For E_2, we get an interesting theorem!

This is automatic theorem discovery using MCCGS.



References

- D. Lazard, F. Rouillier, *Solving parametric polynomial systems*, J. Symbolic Comput. **42** (2007), 636–667.
- M. Manubens, A. Montes, *Minimal canonical comprehensive Groebner systems*, arXiv:0611948 [math.AC].
- A. Montes, T. Recio, *Automatic discovery of geometric theorems using minimal canonical comprehensive Groebner systems*, arXiv:0703483 [math.AG].
- A. Suzuki, Y. Sato, *A simple algorithm to compute comprehensive Gröbner bases*, Proceedings ISSAC 2006, ACM, 326–331.
- M. Wibmer, *Gröbner bases for families of affine or projective schemes*, arXiv:0608019 [math.AC].

Changing Gröbner Bases

- Sometimes one Gröbner basis (say grevlex) is easy to find while another Gröbner basis (say lex) is harder.
- In the 0-dimensional case, one can use the FGLM algorithm of Faugère, Gianni, Lazard and Mora (1993).
- For arbitrary ideals, one can use the Gröbner walk of Collart, Kalkbrener and Mall (1997).
- Avoiding “bad walks” sometimes requires perturbations and large integer arithmetic.
- The generic Gröbner walk of Fukada, Jensen, Lauritzen and Thomas (2007) avoids this problem.

The Gröbner Cone

Definition

Let $(G,>)$ be the reduced Gröbner basis of $I \subseteq k[\mathbf{x}]$. The Gröbner cone $C_>(I) \subseteq \mathbb{R}^n_+$ consists of all $w \in \mathbb{R}^n_+$ such that

$$w\cdot u \geq w\cdot v \quad \text{i.e.,} \quad w\cdot(u-v) \geq 0,$$

where $\mathbf{x}^u = \mathrm{LM}(g)$, $g \in G$, and $\mathbf{x}^v \neq \mathbf{x}^u$ appears in g.

Example

$\{y^3-x^2, x^3-y^2+x\}$ is a reduced Gröbner basis for $I = \langle x^2-y^3, x^3-y^2+x\rangle$ using grlex with $x > y$. Then $C_>(I) \subseteq \mathbb{R}^2_+$ is defined by

$$w\cdot(-2,3) \geq 0, \;\; w\cdot(3,-2) \geq 0, \;\; w\cdot(2,0) \geq 0.$$

The Gröbner Fan

Theorem

Fix an ideal $I \subseteq k[\mathbf{x}]$.

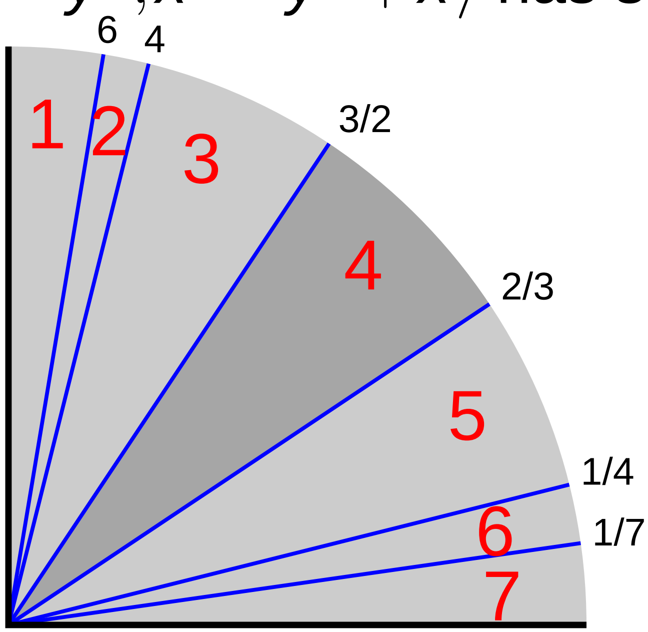
1. *As we vary over all monomial orders on $k[\mathbf{x}]$, I has only finitely many distinct reduced Gröbner bases.*
2. *Two distinct Gröbner cones of I intersect in a common face of each.*

Corollary

The finitely many distinct Gröbner cones form a fan, called the (restricted) Gröbner fan, whose support is the first orthant $\mathbb{R}^n_+$ of $\mathbb{R}^n$.

Example

The Gröbner fan of $\langle x^2 - y^3, x^3 - y^2 + x \rangle$ has seven cones:



1: lex with $y > x$

$\{x^8 - 3x^6 + 3x^4 - x^3 - x^2, xy - x^7 + 2x^5 - x^3 + x^2, y^2 - x^3 - x\}$

4: grlex or grevlex with $y > x$ or $x > y$

$\{y^3 - x^2, x^3 - y^2 + x\}$

7: lex with $x > y$

$\{y^9 - 2y^6 - y^4 + y^3, x - y^7 + y^4 - y^2\}$

The Gröbner Walk

Let $(G, >)$ be a reduced Gröbner basis of I.

Definition

If $g = \sum_v a_v \mathbf{x}^v \in G$ and $w \in \mathbb{R}^n_+$, then

$$\mathrm{in}_w(g) := \sum_{w \cdot v = \max} a_v \mathbf{x}^v.$$

Lemma

w is in the interior of $C_>(I) \iff \mathrm{in}_w(g) = \mathrm{LT}(g)\ \forall g \in G$.

Intuition for the Gröbner Walk

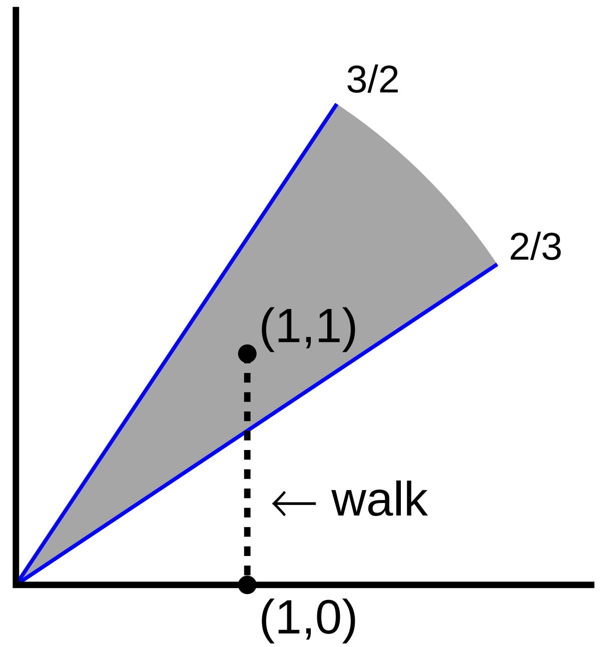
If w lies on the boundary of $C_>(I)$, then:

- $\langle \mathrm{in}_w(G) \rangle$ is "close" to a monomial ideal.
- Finding the Gröbner basis on the other side is "easy".

Example

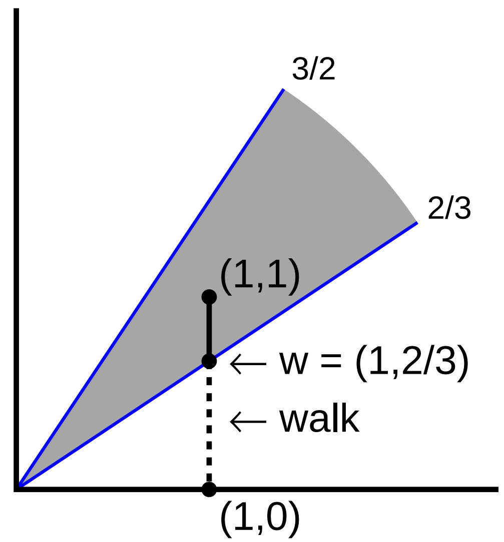
$I = \langle x^2 - y^3, x^3 - y^2 + x \rangle$ has reduced Gröbner basis $G = \{y^3 - x^2, x^3 - y^2 + x\}$ for grevlex with $x > y$. We will "walk" to the reduced Gröbner basis for lex with $x > y$. Note:

- $(1,1)$ is in the interior of the initial Gröbner cone.
- $(1,0)$ is in the target Gröbner cone.



Example

- Compute that we leave the initial cone at $w = (1, 2/3)$.



- Compute $\text{in}_w(G) = \{y^3 - x^2, x^3\}$ and compute a lex Gröbner basis $H = \{x^2 - y^3, xy^3, y^6\}$ of $\langle \text{in}_w(G) \rangle$.

Lifting Lemma

A Gröbner basis (possibly non-reduced) for $>_{w,lex}$ is

$$\{h - h^G \mid h \in H\}.$$

Example

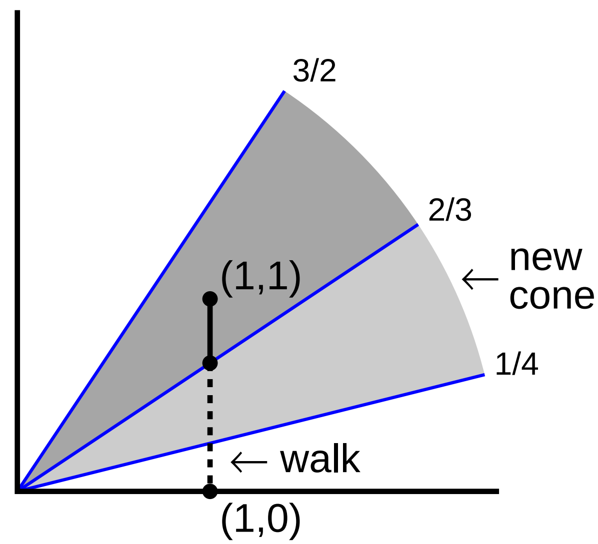
- The lemma gives the Gröbner basis:

$$\{x^2 - y^3, xy^3 - y^2 - x, y^6 - xy^2 - x^2\}.$$

Reducing gives the new reduced Gröbner basis:

$$\{x^2 - y^3, xy^3 - y^2 - x, y^6 - xy^2 - y^3\}$$

and new Gröbner cone:

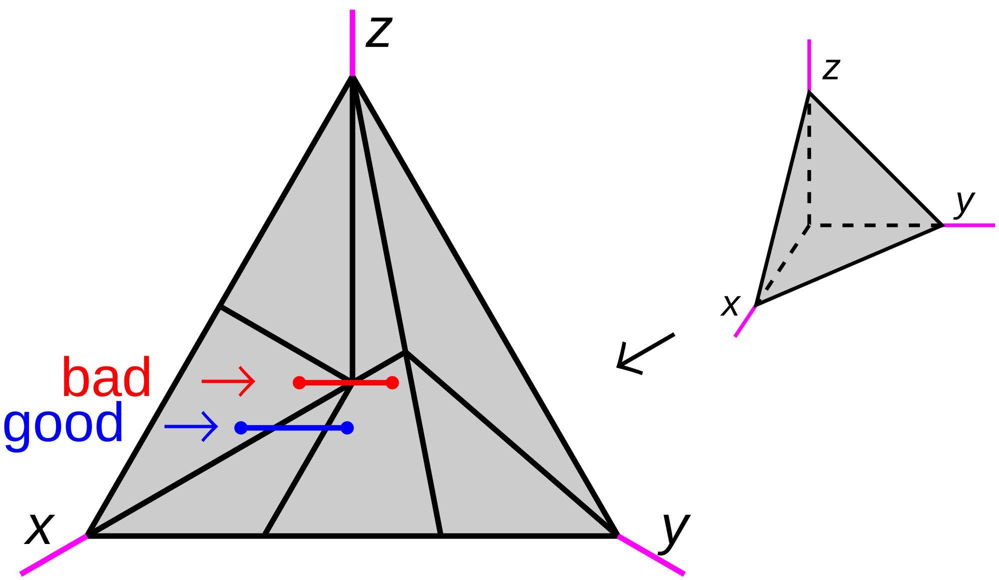


- Repeat, starting with the new cone!

Problems

A walk can be:

- good when it crosses only walls.
- bad when it meets an intersection of walls.



(This is a 2-dimensional slice of a Gröbner fan in $\mathbb{R}^3_+$.)

Perturbations are used to ensure that the walk is good, but this causes problems with w (large integers).

The Generic Gröbner Walk

Goal

Convert a reduced Gröbner basis $(G_1, >_1)$ to $(G_2, >_2)$.

The generic Gröbner walk picks $w_i \in C_{>_i}(I)$ that are sufficiently generic and computes *consequences* of w_i without knowing w_i *explicitly*.

Key Question

When does $w = (1-t)w_1 + t\,w_2$ leave $C_{>_1}(I)$?

$C_{>_1}(I)$ consists of those $w \in \mathbb{R}^n_+$ satisfying

$$w \cdot u \geq w \cdot v, \quad g = a_u \mathbf{x}^u + \cdots + a_v \mathbf{x}^v + \cdots \ \in G_1$$

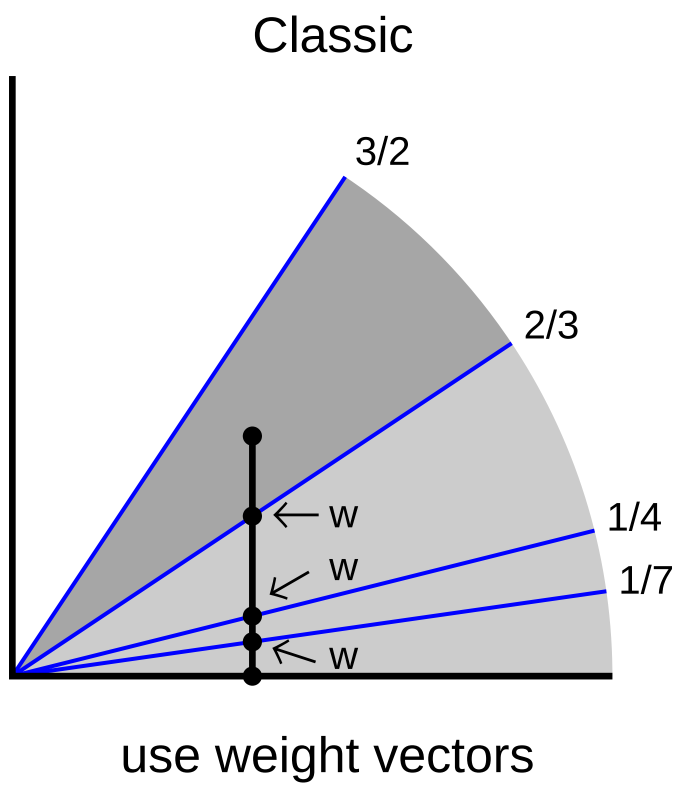
Set $\Delta(G_1) = \{u - v \text{ as above}\}$, so that $C_{>_1}(I)$ is

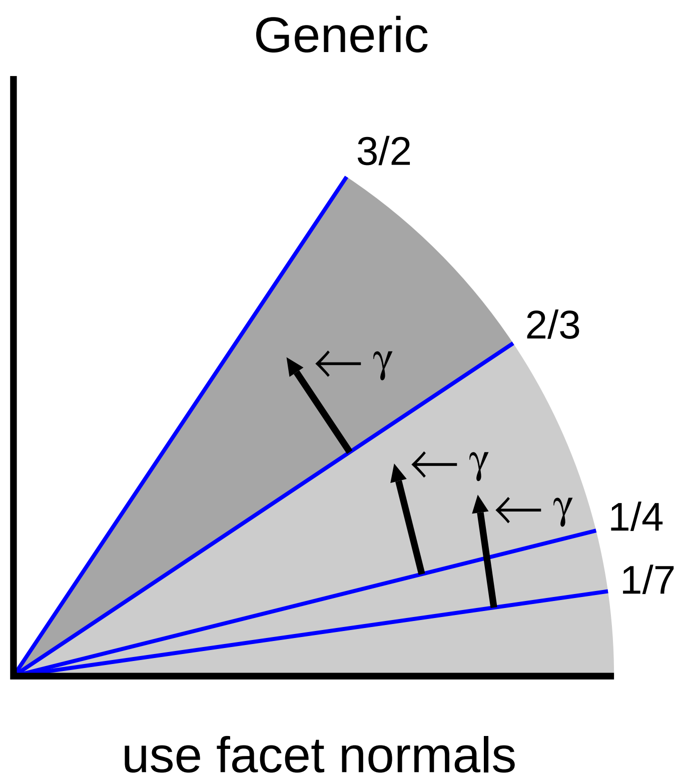
$$w \cdot \gamma \geq 0 \ \ \forall \gamma \in \Delta(G_1), \quad w \in \mathbb{R}^n_+.$$

Elements of $\Delta(G_1)$ are called facet normals.

Normals Instead of Weights

The generic Gröbner walk replaces weight vectors with facet normals to keep track of the walk.





Facet Normals Tell Us Everything

Suppose the classic Gröbner walk crosses cones at a generic point w of the facet defined by γ.

If $g = a_u\mathbf{x}^u + \cdots + a_v\mathbf{x}^v + \cdots \in G$, then:

- $\mathrm{in}_w(g)$ equals $a_u\mathbf{x}^u$ plus those terms $a_v\mathbf{x}^v$ with $w\cdot u = w\cdot v$, i.e., $w\cdot(u-v) = 0$.
- By genericity, this happens $\Leftrightarrow u - v$ is parallel to γ.
- Thus, writing $\mathrm{in}_\gamma(g)$ instead of $\mathrm{in}_w(g)$, we have

$$\mathrm{in}_\gamma(g) = a_u\mathbf{x}^u + \textstyle\sum_{u-v||\gamma} a_v\mathbf{x}^v.$$

Furthermore:

- H is the Gröbner basis of $\langle \mathrm{in}_\gamma(G)\rangle$ for $>_2$.
- The new Gröbner basis is $\{f - f^G \mid f \in H\}$, where f^G is the remainder of f on division by G under $>_1$.

Find the Fact Normal

Lemma

$w = (1-t)w_1 + t\,w_2$ leaves the Gröbner cone of G when

$$t = \min\{t_\gamma \mid \gamma \in \Delta(G),\ \underbrace{w_1\cdot\gamma > 0,\ w_2\cdot\gamma < 0}_{w_1, w_2 \text{ on opposite sides}}\},$$

where

$$t_\gamma := \frac{1}{1 - \dfrac{w_2\cdot\gamma}{w_1\cdot\gamma}}.$$

Any γ with minimal t_γ is the desired facet normal!

Let's compare t_γ and t_δ. Since w_2 is generic for $>_2$,

$$t_\gamma \le t_\delta \Leftrightarrow w_2\cdot(w_1\cdot\delta)\gamma \le w_2\cdot(w_1\cdot\gamma)\delta \Leftrightarrow (w_1\cdot\delta)\gamma \le_2 (w_1\cdot\gamma)\delta.$$

This removes w_2 from the picture.

Find the Facet Normal

To remove w_1, represent $>_2$ using $\tau_1, \ldots, \tau_n$. This means

$$u <_2 v \Leftrightarrow \tau_1 \cdot u < \tau_1 \cdot v, \text{ or } \tau_1 \cdot u = \tau_1 \cdot v \text{ and } \tau_2 \cdot u < \tau_2 \cdot v, \text{ or } \ldots$$

Then testing $(w_1 \cdot \delta)\gamma \leq_2 (w_1 \cdot \gamma)\delta$ requires checking

$$\tau_i \cdot \big((w_1 \cdot \delta)\gamma\big) \leq \tau_i \cdot \big((w_1 \cdot \gamma)\delta\big).$$

Rewrite this as

$$w_1 \cdot \big((\tau_i \cdot \gamma)\delta\big) \leq w_1 \cdot \big((\tau_i \cdot \delta)\gamma\big).$$

Since w_1 is generic for $>_1$, this is equivalent to

$$(\tau_i \cdot \gamma)\delta \leq_1 (\tau_i \cdot \delta)\gamma.$$

This removes w_1 from the picture.

Example

As before, we convert a reduced grevlex Gröbner basis with $x > y$ for $I = \langle x^2 - y^3, x^3 - y^2 + x \rangle$ to lex with $x > y$.

We assume we have done one step to obtain the first facet normal $\gamma = (-2, 3)$ and the reduced Gröbner basis

$$G = \{x^2 - y^3, xy^3 - y^2 - x, y^6 - xy^2 - y^3\}$$

This gives $\Delta(G) = \{(2,-3),(1,1),(0,3),(-1,4)\}$.

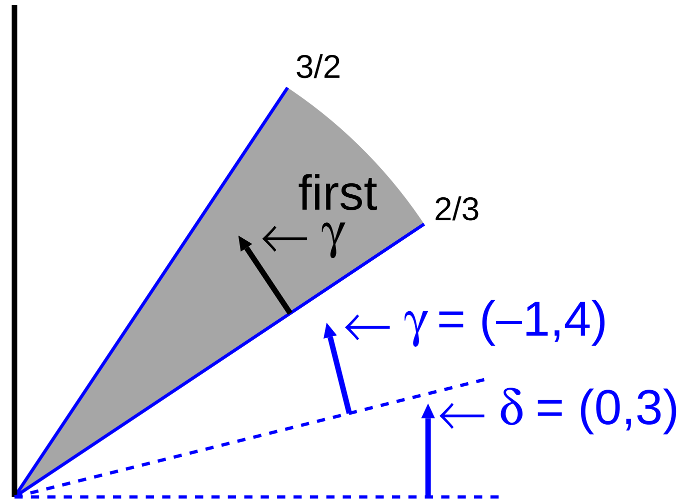
The second facet normal lies in

$$\{\gamma \in \Delta(G) \mid \underbrace{\gamma >_1 0,\ \gamma <_2 0}_{\gamma \text{ separates } >_1, >_2}\} = \{\underbrace{(-1,4)}_{\gamma}, \underbrace{(0,3)}_{\delta}\}$$

Which do we use: γ or δ?

Example

Here is the picture:



Then:

- $>_1$ is grevlex with $x > y$ and $>_2$ is lex with $x > y$.
- Represent $>_2$ using:

$$\tau_1 = (1,0), \tau_2 = (0,1).$$

- Our previous analysis gives:

$$t_\gamma < t_\delta \Leftrightarrow (\tau_1 \cdot \gamma)\delta <_1 (\tau_1 \cdot \delta)\gamma, \text{ or } (\tau_1 \cdot \gamma)\delta = (\tau_1 \cdot \delta)\gamma \text{ and } \ldots$$
$$\Leftrightarrow (0,-3) <_1 (0,0).$$

Conclusion: The second γ is $(-1,4)$.

Summary

The Generic Gröbner Walk

- Input: A reduced Gröbner basis $(G_1, >_1)$ of I.
- Output: A reduced Gröbner basis $(G_2, >_2)$ of I.
- 1: Use G_1 to find the first γ using $(\tau_i \cdot \gamma)\delta \leq_1 (\tau_i \cdot \delta)\gamma$.
- 2: Use γ to find $\mathrm{in}_\gamma(G_1)$.
- 3: Find a reduced Gröbner basis $(H, >_2)$ of $\langle \mathrm{in}_\gamma(G_1) \rangle$.
- 4: Lift to $H' = \{f - f^G \mid f \in H\}$.
- 5: Autoreduce to get the next reduced Gröbner basis.
- 6: Iterate!

Final Comments

Another interesting variation on the Gröbner walk has been proposed by Tran (2007). It uses the following definition.

Definition

Given $I \subseteq k[\mathbf{x},\mathbf{y}]$, a monomial order $>$ is ideal-specific for I for eliminating $\mathbf{x}$ if the reduced Gröbner basis $(G,>)$ has the property that for any $g \in G$,

$$\mathrm{LT}(g) \in k[\mathbf{y}] \Rightarrow g \in k[\mathbf{y}].$$

Lemma

Given $I \subseteq k[\mathbf{x},\mathbf{y}]$ and $(G,>)$ as in the definition, $G \cap k[\mathbf{y}]$ is a Gröbner basis for the elimination ideal $I \cap k[\mathbf{y}]$.

Final Comments

Lemma

Suppose a Gröbner cone $C_>(I)$ contains a weight vector

$$w = (\underbrace{\beta_1, \ldots, \beta_s}_{\mathbf{x} \textit{ variables}}, \underbrace{0, \ldots, 0}_{\mathbf{y} \textit{ variables}}).$$

Then $>$ is ideal-specific for I for eliminating **x**.

Tran has applied this to elimination theory via the classic Gröbner walk (with perturbations introduced in 2000).

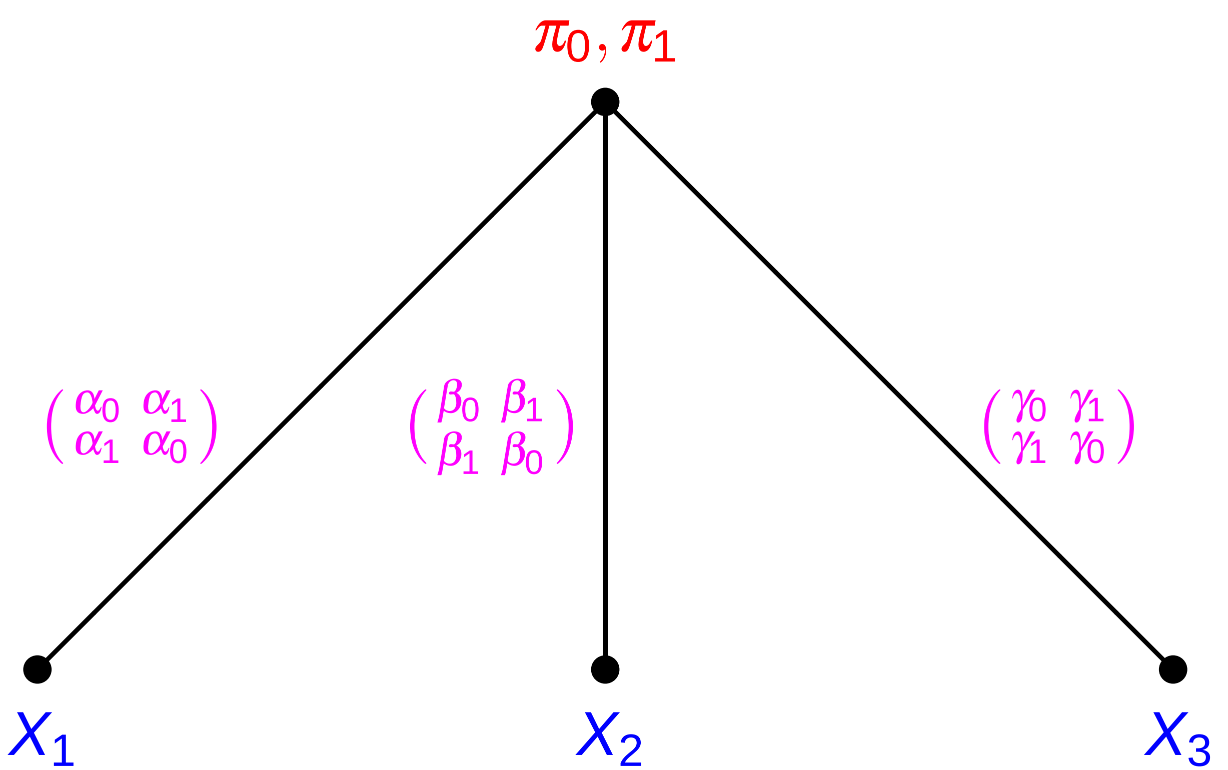
- The target is an elimination order $>_2$, but the walk stops as soon as if finds an ideal-specific elimination order.
- In examples, the Gröbner walk traverses 92 cones, but finds an ideal-specific elimination order after 86.
- Since the later cones can require the most computation, this can cut the running time in half.

References

- S. Collart, M. KalkBrener, D. Mall, *Converting bases with the Gröbner walk*, J. Symbolic Comput. **24** (1997), 465–469.
- D. Cox, J. Little, D. O'Shea, *Using Algebraic Geometry*, Second Edition, Graduate Texts in Mathematics, Springer, New York, 2005.
- K. Fukada, A. N. Jensen, N. Lauritzen, R. Thomas, *The generic Gröbner walk*, J. Symbolic Comput. **42** (2007), 298–312.
- Q.-N. Tran, *A new class of term orders for elimination*, J. Symbolic Comput. **42** (2007), 533–548.

Phylogenetic Trees

Consider a rooted tree with a probability distribution at the root, a binary random variable at each leaf, and a transition matrix along each edge.



This is the Cantor-Jukes binary model for $K_{1,3}$.

The Probability Distribution

The matrices $A^1 = \begin{pmatrix} \alpha_0 & \alpha_1 \\ \alpha_1 & \alpha_0 \end{pmatrix}$, $A^2 = \begin{pmatrix} \beta_0 & \beta_1 \\ \beta_1 & \beta_0 \end{pmatrix}$, $A^3 = \begin{pmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{pmatrix}$ give

$$P(X_\ell = j) = \pi_0 A^\ell_{0j} + \pi_1 A^\ell_{1j}.$$

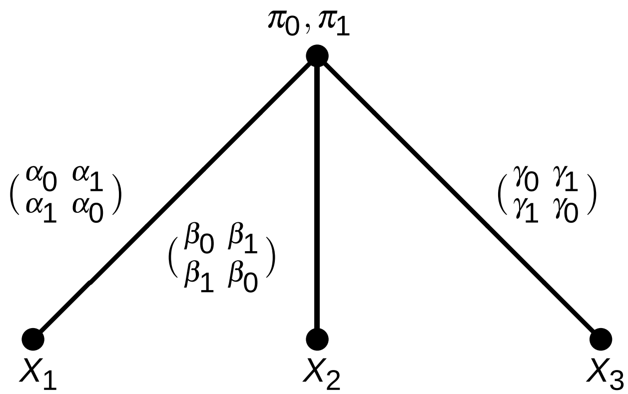
Experimentally, we can measure the probabilities

$$p_{ijk} := P(X_1 = i, X_2 = j, X_3 = k).$$

The goal is to determine $\pi_i, \alpha_i, \beta_i, \gamma_i$. Note that

$$p_{000} = \pi_0\alpha_0\beta_0\gamma_0 + \pi_1\alpha_1\beta_1\gamma_1$$
$$p_{001} = \pi_0\alpha_0\beta_0\gamma_1 + \pi_1\alpha_1\beta_1\gamma_0,$$

and so on.



Phylogenetic Invariants

The probabilities p_{ijk} are not independent. For example, regardless of how we assign $\pi_i, \alpha_i, \beta_i, \gamma_i$, the p_{ijk} always satisfy the relation

$$\begin{aligned} &p_{001}p_{010} + p_{001}p_{100} - p_{000}p_{011} - p_{000}p_{101} \\ &+ p_{100}p_{111} - p_{101}p_{110} + p_{010}p_{111} - p_{011}p_{110} = 0. \end{aligned}$$

This is an example of a binary phylogenetic invariant.

Binary phylogenetic invariants form an ideal in $k[p_{ijk}]$. Note:

- We can compute this ideal by eliminating $\pi_i, \alpha_i, \beta_i, \gamma_i$ from the equations defining the p_{ijk}.
- Knowing the phylogenetic invariants is useful since they mean we don't have to find all p_{ijk}—we can find some and solve the above equations for the rest.
- Once we have the p_{ijk}, we can solve for $\pi_i, \alpha_i, \beta_i, \gamma_i$.

Change Coordinates

A Fourier transform gives a linear change of coordinates that transforms both p_{ijk} and the parameters $\pi_i, \alpha_i, \beta_i, \gamma_i$ into new variables q_{ijk} and new parameters r_i, a_i, b_i, c_i such that

$$q_{000} = r_0 a_0 b_0 c_0,\ q_{001} = r_1 a_0 b_0 c_1,\ \ldots$$

Eliminating the parameters gives the phylogenetic ideal

$$\langle q_{001}q_{110} - q_{000}q_{111}, q_{010}q_{101} - q_{000}q_{111}, q_{100}q_{011} - q_{000}q_{111}\rangle.$$

This is a example of a toric ideal (a prime ideal generated by differences of monomials). Ordering the variables

$$q_{000} > q_{001} > q_{010} > q_{011} > q_{100} > q_{101} > q_{110} > q_{111},$$

(the reverse of their binary values), a lex Gröbner basis is

$$\{q_{000}q_{111} - q_{011}q_{100}, q_{001}q_{110} - q_{011}q_{100}, q_{010}q_{101} - q_{011}q_{100}\}.$$

Some Results

Theorem (Sturmfels and Sullivant (2005))

The computation of the binary phylogenetic ideal can be reduced to the case of $K_{1,n}$.

For $K_{1,n}$, we have:

- 2^n variables $q_{i_1\ldots i_n}$, $i_j = 0,1$.
- $2(n+1)$ parameters $a_i^{(j)}$, $i = 0,1$, $j = 0,1,\ldots,n$.
- 2^n equations $q_{i_1\ldots i_n} = a^{(0)}_{i_1+\cdots+i_n} a^{(1)}_{i_1} \cdots a^{(n)}_{i_n}$.

Theorem (Sturmfels and Sullivant (2005))

The binary phylogenetic ideal $I_n \subseteq k[q_{i_1\ldots i_n}]$ obtained by eliminating parameters is generated by degree 2 binomials.

A Gröbner basis?

The number of degree 2 generators of I_n increases with n:

n	3	4	5	6
# min gens	3	30	195	1050

For $n = 3$, the degree 2 generators form a Gröbner basis.

Question

Is this true in general? In other words, does I_n always have a degree 2 Gröbner basis?

Recent work of Chifman and Petrović (2007) says that the answer is yes.

Set-up for the Gröbner Basis

The degree 2 Gröbner basis G_n of I_n is described recursively.

- $R_n := k[q_{i_1 \dots i_n}]$ has lex order with
$$q_{0\dots 00} > q_{0\dots 01} > q_{0\dots 10} > q_{0\dots 11} > \cdots > q_{1\dots 10} > q_{1\dots 11}.$$
- For $1 \le j \le n$, define $\pi_j : R_n \to R_{n-1}$ by
$$\pi_j(q_{i_1\dots i_n}) = q_{i_1\dots \widehat{i_j} \dots i_n}.$$
- For $A, B, C, D \in \{0,1\}^n$ with $q_A > q_B, q_A > q_C > q_D$, set
$$g = q_A q_B - q_C q_D.$$

Goal

Given G_{n-1}, describe which g's lie in G_n.

Describe the Gröbner Basis

Recall that G_3 is

$$\{q_{000}q_{111} - q_{011}q_{100}, q_{001}q_{110} - q_{011}q_{100}, q_{010}q_{101} - q_{011}q_{100}\}.$$

Definition

Assume G_{n-1} has been defined for some $n \geq 4$. Then G_n consists of the following two types of $g = q_A q_B - q_C q_D$:

- Type 1: For some j, we have $A_j = B_j = C_j = D_j$ and $\pi_j(g) \in G_{n-1}$.
- Type 2: For all j, we have $A_j + B_j = C_j + D_j = 1$ and $\pi_j(g) \in G_{n-1}$.

Theorem (Chifman and Petrović (2007))

G_n is a lex Gröbner basis of I_n.

Consequences

Corollary

For any tree, the binary phylogenetic ideal has a Gröbner basis consisting of degree 2 binomials.

Corollary

When we regard the field k as graded module over the quotient ring R_n/I_n, its free resolution is

$$\cdots \to (R_n/I_n)^{\beta_2}(-2) \to (R_n/I_n)^{\beta_1}(-1) \to R_n/I_n \to k \to 0.$$

It follows that R_n/I_n is a Koszul algebra.

This corollary is a good example of how Gröbner bases can be used to prove theoretical results in commutative algebra and algebraic geometry.

A Final Question

Question

Sequence A032263 from Sloane's On-Line Encyclopedia of Integer Sequences begins

$$0,\ 0,\ 0,\ 3,\ 30,\ 195,\ 1050,\ 5103,\ 23310,\ \ldots$$

This gives the number of 2-element proper antichains in an n-element set. The formula for the nth term is

$$\tfrac{1}{2}(4^n - 3\cdot 3^n + 3\cdot 2^n - 1).$$

Is this $|G_n|$?

References

- J. A. Cavender, J. Felsenstein, *Invariants of phylogenies: a simple case with discrete states*, J. Classification **4** (1987), 57–71.
- J. Chifman, S. Petrović, *Toric ideals of phylogenetic invariants for the general group-based model on claw trees* $K_{1,n}$, arXiv:0702368 [math.AC].
- S. Evans, T. Speed, *Invariants of some probability models used in phylogenetic inference*, Ann. Statist. **21** (1993), 355–377.
- L. Pachter, B. Sturmfels, *Algebraic Statistics for Computational Biology*, Cambridge University Press, New York, 2005.
- B. Sturmfels, S. Sullivant, *Toric ideals of phylogenetic invariants*, J. Comp. Biol. **12** (2005), 204–228.