Lecture 9: Elimination Ideals and Resultants

Rafael Oliveira

University of Waterloo
Cheriton School of Computer Science

rafael.oliveira.teaching@gmail.com

February 8, 2021

Overview

- Solving Polynomial Equations
  - Elimination Theorem
- Extension Theorem
  - Resultants
- Conclusion
- Acknowledgements

Solving Polynomial Equations

- Last lecture we saw how to generalize division algorithm and Gaussian Elimination

Solving Polynomial Equations

- Last lecture we saw how to generalize division algorithm and Gaussian Elimination
- Groebner bases were crucial to make our generalized division algorithm work

Solving Polynomial Equations

- Last lecture we saw how to generalize division algorithm and Gaussian Elimination
- Groebner bases were crucial to make our generalized division algorithm work
- How can we use Groebner bases to solve polynomial equations? After all, Gaussian Elimination helps us solve linear systems of equations

Solving Polynomial Equations

- Last lecture we saw how to generalize division algorithm and Gaussian Elimination
- Groebner bases were crucial to make our generalized division algorithm work
- How can we use Groebner bases to solve polynomial equations? After all, Gaussian Elimination helps us solve linear systems of equations
- Today we will learn:
  1. *Elimination Theorem*: how to "eliminate" variables from our system of polynomial equations
  2. *Extension Theorem*: how to "extend" partial solutions to complete solutions

Elimination Theorem

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

Elimination Theorem

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

- $I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$. Want $V(I)$.

Elimination Theorem

$x > y > z$

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

- $I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$. Want $V(I)$.
- Computing Groebner basis of I with respect to lex order:

$$G = (\underbrace{x+y+z^2-1}_{x,y,z}, \underbrace{y^2-y-z^2+z}_{y,z}, \underbrace{2yz^2+z^4-z^2}_{y,z}, \underbrace{z^6-4z^4+4z^3-z^2}_{z})$$

Elimination Theorem

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

- $I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$. Want $V(I)$.
- Computing Groebner basis of I with respect to lex order:

$$G = (x + y + z^2 - 1, y^2 - y - z^2 + z, 2yz^2 + z^4 - z^2, z^6 - 4z^4 + 4z^3 - z^2)$$

- Since $G = I$ we know both systems have same zero set! What is special about the Groebner basis set of equations?

in lex order!

Elimination Theorem

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

- $I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$. Want $V(I)$.
- Computing Groebner basis of I with respect to lex order:

$$G = (x+y+z^2-1, y^2-y-z^2+z, 2yz^2+z^4-z^2, z^6-4z^4+4z^3-z^2)$$

- Since $G = I$ we know both systems have same zero set! What is special about the Groebner basis set of equations?
- Last polynomial only depends on z *elimination step*

Elimination Theorem

- Example:

$$x^2 + y + z = 1$$
$$x + y^2 + z = 1$$
$$x + y + z^2 = 1$$

- $I = (x^2 + y + z - 1, x + y^2 + z - 1, x + y + z^2 - 1)$. Want $V(I)$.
- Computing Groebner basis of I with respect to lex order:

$$G = (x + y + z^2 - 1, y^2 - y - z^2 + z, 2yz^2 + z^4 - z^2, z^6 - 4z^4 + 4z^3 - z^2)$$

- Since $G = I$ we know both systems have same zero set! What is special about the Groebner basis set of equations?
- Last polynomial only depends on z *elimination step*
- Can find all possible z's and propagate it up to find y and then x *extension step*

Elimination Theorem

- Main idea of elimination theory is to find the part of the ideal that “depends on less variables”

Elimination Theorem

- Main idea of elimination theory is to find the part of the ideal that "depends on less variables"
- Given $I \subset \mathbb{F}[x_1, \ldots, x_n]$, the ℓ^{th} elimination ideal I_ℓ is the ideal of $\mathbb{F}[x_{\ell+1}, \ldots, x_n]$ given by:

$$I_\ell := I \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$$

only polynomials from I that only depend on n-ℓ last variables.

Elimination Theorem

- Main idea of elimination theory is to find the part of the ideal that "depends on less variables"
- Given $I \subset \mathbb{F}[x_1, \ldots, x_n]$, the ℓ^{th} elimination ideal I_ℓ is the ideal of $\mathbb{F}[x_{\ell+1}, \ldots, x_n]$ given by:

$$I_\ell := I \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$$

- The *elimination step* is to find these ideals I_ℓ for all $\ell \in [n]$.

Elimination Theorem

- Main idea of elimination theory is to find the part of the ideal that "depends on less variables"
- Given $I \subset \mathbb{F}[x_1, \ldots, x_n]$, the ℓ^{th} elimination ideal I_ℓ is the ideal of $\mathbb{F}[x_{\ell+1}, \ldots, x_n]$ given by:

$$I_\ell := I \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$$

- The *elimination step* is to find these ideals I_ℓ for all $\ell \in [n]$.
- *Elimination Theorem*

  For any ideal $I \subset \mathbb{F}[x_1, \ldots, x_n]$, if G is a Groebner basis of I with respect to the *lexicographic order* $x_1 \succ x_2 \succ \ldots \succ x_n$, then

$$G_\ell := G \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$$

  is a Groebner basis of I_ℓ.

Proof of Elimination Theorem

- Suffices to show that $LM(I_\ell) = LM(G_\ell)$

$G = \{g_1, \ldots, g_t\}$

$f \in I_\ell = I \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$

$\Rightarrow LM(f) \in \mathbb{F}[x_{\ell+1}, \ldots, x_n]$

$\Rightarrow$ (lex order) $f = \sum_{i=1}^{t} B_i g_i$

$\hookrightarrow B_i g_i \neq 0$ only if $g_i \in I_\ell$

division algorithm

multideg$(f) \geq$ multideg$(B_i g_i)$

Proof of Elimination Theorem

- Suffices to show that $LM(I_\ell) = LM(G_\ell)$
- So in our example above, the last polynomial was *the best way* to eliminate variables x, y from our system.

- Solving Polynomial Equations
  - Elimination Theorem
- Extension Theorem
  - Resultants
- Conclusion
- Acknowledgements

Extension Theorem

- Now that we know how eliminate variables from our system of polynomial equations, we need to learn how to reconstruct a full solution based on our partial solution

Extension Theorem

- Now that we know how eliminate variables from our system of polynomial equations, we need to learn how to reconstruct a full solution based on our partial solution
- Given solution $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell) \subseteq \mathbb{F}^{n-\ell}$ we want to find a solution $(a_\ell, \ldots, a_n) \in V(I_{\ell-1}) \subseteq \mathbb{F}^{n-\ell+1}$

Extension Theorem

- Now that we know how eliminate variables from our system of polynomial equations, we need to learn how to reconstruct a full solution based on our partial solution
- Given solution $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell) \subseteq \mathbb{F}^{n-\ell}$ we want to find a solution $(a_\ell, \ldots, a_n) \in V(I_{\ell-1}) \subseteq \mathbb{F}^{n-\ell+1}$
- So we are essentially trying to solve a system of *univariate polynomials*

$I_{\ell-1} \subseteq \mathbb{F}[x_\ell, \ldots, x_n] \leftarrow x_\ell \ldots$

$I_\ell = I_{\ell-1} \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n] \leftarrow$ don't depend on x_ℓ

Extension Theorem

- Now that we know how eliminate variables from our system of polynomial equations, we need to learn how to reconstruct a full solution based on our partial solution
- Given solution $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell) \subseteq \mathbb{F}^{n-\ell}$ we want to find a solution $(a_\ell, \ldots, a_n) \in V(I_{\ell-1}) \subseteq \mathbb{F}^{n-\ell+1}$
- So we are essentially trying to solve a system of *univariate polynomials*
- What could go wrong? Partial solutions that don't extend to complete solutions. Example:

$$xy = 1, \quad xz = 1 \quad \text{partial solution } y = z = 0$$

Groebner basis: $(xy - 1, xz - 1, y - z)$

Extension Theorem

- Now that we know how eliminate variables from our system of polynomial equations, we need to learn how to reconstruct a full solution based on our partial solution
- Given solution $(a_{\ell+1}, \ldots, a_n) \in V(I_\ell) \subseteq \mathbb{F}^{n-\ell}$ we want to find a solution $(a_\ell, \ldots, a_n) \in V(I_{\ell-1}) \subseteq \mathbb{F}^{n-\ell+1}$
- So we are essentially trying to solve a system of *univariate polynomials*
- What could go wrong? Partial solutions that don't extend to complete solutions. Example:

$$xy = 1, \quad xz = 1 \qquad \text{partial solution } y = z = 0$$

  Groebner basis: $(xy - 1, xz - 1, y - z)$
- Extension theorem gives us a sufficient condition to extend partial solutions.

Extension Theorem

lexicographic order

- *Extension Theorem*

  Let $\mathbb{F}$ be an *algebraically closed* field, $I := (f_1, \ldots, f_s) \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and let I_1 be the first elimination ideal of I. For each $1 \leq i \leq s$, write f_i as

$$f_i = c_i(x_2, \ldots, x_n) \cdot x_1^{d_i} + \text{ lower degree terms in } x_1$$

where c_i's are non-zero and $d_i \geq 0$. If

$$(a_2, \ldots, a_n) \in V(I_1)$$

that is, it is a partial solution, and if

$$(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$$

not in zero set of leading coefficients

then there is $a_1 \in \mathbb{F}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.

then we can extend to full solution

Extension Theorem

- *Extension Theorem*

  Let $\mathbb{F}$ be an *algebraically closed* field, $I := (f_1, \ldots, f_s) \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and let I_1 be the first elimination ideal of I. For each $1 \leq i \leq s$, write f_i as

  $$f_i = c_i(x_2, \ldots, x_n) \cdot x_1^{d_i} + \text{ lower degree terms in } x_1$$

  where c_i's are non-zero and $d_i \geq 0$. If

  $$(a_2, \ldots, a_n) \in V(I_1)$$

  that is, it is a partial solution, and if

  $$(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$$

  then there is $a_1 \in \mathbb{F}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.
- Extension step fails then the leading coefficients must vanish

Proof of Extension Theorem

- Let $G = (g_1, \ldots, g_t)$ be a Groebner basis of $I \subseteq \mathbb{F}[x_1, \ldots, x_n]$ with respect to the lex order. For $1 \leq j \leq t$, let

$$g_j = c_j(x_2, \ldots, x_n) \cdot x_1^{d_j} + \quad \text{lower degree terms in } x_1$$

where $d_j \geq 0$ and $c_j \in \mathbb{F}[x_2, \ldots, x_n]$ is non-zero.
Let $\mathbf{a} \in V(I_1) \subseteq \mathbb{F}^{n-1}$ be a partial solution such that $\mathbf{a} \notin V(c_1, \ldots, c_t)$. Then

generated by poly in Gröbner basis

$$I_{\mathbf{a}} := \{f(x_1, \mathbf{a}) \mid f \in I\} = (g_o(x_1, \mathbf{a})) \subseteq \mathbb{F}[x_1]$$

where $g_o \in G$ satisfies $c_o(\mathbf{a}) \neq 0$ and g_o has minimal x_1 degree among all elements $g_j \in G$ with $c_j(\mathbf{a}) \neq 0$. Moreover

1. $\deg(g_o(x_1, \mathbf{a})) > 0$
2. If $g_o(a_1, \mathbf{a}) = 0$ for $a_1 \in \mathbb{F}$, then $(a_1, \mathbf{a}) \in V(I)$

$\bar{a} \notin V(\hat{c}_1, \ldots, \hat{c}_s) = V(c_1, \ldots, c_t)$

Proof of Extension Theorem

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).

Proof of Extension Theorem

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).
- Note that $d_o > 0$, otherwise we would have $g_o = c_o$, which would imply $g_o(x_1, \mathbf{a}) = c_o(\mathbf{a}) \neq 0$, which implies $\mathbf{a} \notin I_1$

Proof of Extension Theorem

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).
- Note that $d_o > 0$, otherwise we would have $g_o = c_o$, which would imply $g_o(x_1, \mathbf{a}) = c_o(\mathbf{a}) \neq 0$, which implies $\mathbf{a} \notin I_1$
- We now need to prove that ~~$g_o(x_1)$~~ generates the ideal $I_{\mathbf{a}}$

  $g_o(x_1, \bar{a})$

Proof of Extension Theorem

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).
- Note that $d_o > 0$, otherwise we would have $g_o = c_o$, which would imply $g_o(x_1, \mathbf{a}) = c_o(\mathbf{a}) \neq 0$, which implies $\mathbf{a} \notin I_1$
- We now need to prove that $g_o(x_1)$ generates the ideal $I_{\mathbf{a}}$
- Since $I \subseteq G$ it is enough to show that

$$g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a})) \quad \forall g_j \in G$$

g_j

Proof of Extension Theorem

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).
- Note that $d_o > 0$, otherwise we would have $g_o = c_o$, which would imply $g_o(x_1, \mathbf{a}) = c_o(\mathbf{a}) \neq 0$, which implies $\mathbf{a} \notin I_1$
- We now need to prove that $g_o(x_1)$ generates the ideal $I_\mathbf{a}$
- Since $I \subseteq G$ it is enough to show that

$$g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a})) \quad \forall c_j \in G$$

- We will prove this by induction on the x_1-degree of the g_j's

Proof of Extension Theorem

$d_j = \deg_1(g_j(x_1, \bar{x}))$

- Choose an $g_o \in G$ as in previous slide (minimal x_1-degree among elements of G with non-zero leading term $c_j(\mathbf{a}) \neq 0$).
- Note that $d_o > 0$, otherwise we would have $g_o = c_o$, which would imply $g_o(x_1, \mathbf{a}) = c_o(\mathbf{a}) \neq 0$, which implies $\mathbf{a} \notin I_1$
- We now need to prove that $g_o(x_1)$ generates the ideal $I_\mathbf{a}$
- Since $I \subseteq G$ it is enough to show that

$$g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a})) \quad \forall c_j \in G$$

- We will prove this by induction on the x_1-degree of the c_j's
- Our choice of g_o tells us that $d_o = \deg(g_o(x_1, \mathbf{a}))$. By minimality of d_o, if any g_j is such that

$$\deg(g_j(x_1, \mathbf{a})) < d_o$$

it must have been that $c_j(\mathbf{a}) = 0$. That is, g_j dropped degree on evaluation.

Proof of Extension Theorem

- If there is $g_j \in G$ with $d_j < d_o$ such that $g_j(x_1, \mathbf{a}) \neq 0$, let g_b be the one which *minimizes* the drop in degree when evaluated at $\mathbf{a}$.
- Let $\delta = d_b - \deg(g_b(x_1, \mathbf{a}))$.

new degree

$\deg_1(g_b(x_1, \bar{x}))$

Proof of Extension Theorem

- If there is $g_j \in G$ with $d_j < d_o$ such that $g_j(x_1, \mathbf{a}) \neq 0$, let g_b be the one which *minimizes* the drop in degree when evaluated at $\mathbf{a}$.
- Let $\delta = d_b - \deg(g_b(x_1, \mathbf{a}))$.
- Let

$$S := S(g_o, g_b) = c_o x_1^{d_o - d_b} g_b - c_b g_o$$

$$\deg_1(g_o) < \deg_1(g_b)$$

Proof of Extension Theorem

- If there is $g_j \in G$ with $d_j < d_o$ such that $g_j(x_1, \mathbf{a}) \neq 0$, let g_b be the one which *minimizes* the drop in degree when evaluated at $\mathbf{a}$.
- Let $\delta = d_b - \deg(g_b(x_1, \mathbf{a}))$.

  $c_b(\bar{a}) = 0$ (because g_b drops degree)
- Let

$$S := S(g_o, g_b) = c_o x_1^{d_o - d_b} g_b - c_b g_o$$

- Note that

$$S(x_1, \mathbf{a}) = c_o(\mathbf{a}) x^{d_o - d_b} g_b(x_1, \mathbf{a}) \quad \neq 0$$

  so $\deg(S(x_1, \mathbf{a})) = d_o - d_b + (d_b - \delta) = d_o - \delta$

Proof of Extension Theorem

- If there is $g_j \in G$ with $d_j < d_o$ such that $g_j(x_1, \mathbf{a}) \neq 0$, let g_b be the one which *minimizes* the drop in degree when evaluated at $\mathbf{a}$.
- Let $\delta = d_b - \deg(g_b(x_1, \mathbf{a}))$.
- Let

$$S := S(g_o, g_b) = c_o x_1^{d_o - d_b} g_b - c_b g_o$$

- Note that

$$S(x_1, \mathbf{a}) = c_o(\mathbf{a}) x^{d_o - d_b} g_b(x_1, \mathbf{a})$$

  so $\deg(S(x_1, \mathbf{a})) = d_o - d_b + (d_b - \delta) = d_o - \delta$

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

in lex order

standard rep. in lex order

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

S polynomial of g_o, g_b

  when $B_j g_j \neq 0$.

Proof of Extension Theorem

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

when $B_j g_j \neq 0$.

Proof of Extension Theorem

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

  when $B_j g_j \neq 0$.
- So if g_j appears in standard representation, then $\deg_1(g_j) < d_o$ which implies g_j must *drop degree* or *go to zero* when evaluated at $\mathbf{a}$

Proof of Extension Theorem

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

when $B_j g_j \neq 0$.

- So if g_j appears in standard representation, then $\deg_1(g_j) < d_o$ which implies g_j must *drop degree* or *go to zero* when evaluated at $\mathbf{a}$
- Thus, we have:

$$\deg(B_j(x_1, \mathbf{a})) + \underline{\deg(g_j(x_1, \mathbf{a}))} \leq \underline{\deg_1(B_j)} + \underbrace{\deg_1(g_j) - \delta}_{\text{by minimality of } \delta} < \underline{d_o - \delta}$$

Proof of Extension Theorem

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

  when $B_j g_j \neq 0$.
- So if g_j appears in standard representation, then $\deg_1(g_j) < d_o$ which implies g_j must *drop degree* or *go to zero* when evaluated at $\mathbf{a}$
- Thus, we have:

$$\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a})) \leq \deg_1(B_j) + \deg_1(g_j) - \delta < d_o - \delta$$

- Thus:

$$\deg(S(x_1, \mathbf{a})) \leq \max\{\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a}))\} < d_o - \delta$$

  contradiction.

Proof of Extension Theorem

- Since G is a Groebner basis, $S = \sum_{i=1}^{t} B_j g_j$ standard representation, which implies

$$\deg_1(B_j) + \deg_1(g_j) = \deg_1(B_j g_j) \leq \deg_1(S) < d_o$$

when $B_j g_j \neq 0$.
- So if g_j appears in standard representation, then $\deg_1(g_j) < d_o$ which implies g_j must *drop degree* or *go to zero* when evaluated at $\mathbf{a}$
- Thus, we have:

$$\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a})) \leq \deg_1(B_j) + \deg_1(g_j) - \delta < d_o - \delta$$

- Thus:

$$\deg(S(x_1, \mathbf{a})) \leq \max\{\deg(B_j(x_1, \mathbf{a})) + \deg(g_j(x_1, \mathbf{a}))\} < d_o - \delta$$

contradiction.
- Thus, if g_j dropped degree and it is non-zero after evaluation, it must be $d_j \geq d_o$.

Proof of Extension Theorem

- Now we are ready to prove that $I_a = (g_o(x_1, \mathbf{a}))$ by induction.

Proof of Extension Theorem

- Now we are ready to prove that $I_a = (g_o(x_1, \mathbf{a}))$ by induction.
- By the above, claim is true for any $g_j \in G$ with $d_j < d_o$.
- Let $d \geq d_o$ and assume claim is true for any $g_j \in G$ with $d_j < d$.

Proof of Extension Theorem

- Now we are ready to prove that $I_a = (g_o(x_1, \mathbf{a}))$ by induction.
- By the above, claim is true for any $g_j \in G$ with $d_j < d_o$.
- Let $d \geq d_o$ and assume claim is true for any $g_j \in G$ with $d_j < d$.
- Let $g_i \in G$ be such that $d_i = d$.
- Taking standard representation of $S(g_i, g_o) = \sum_{k=1}^{t} B_k g_k$, where

$$S := c_o g_j - c_j x_1^{d-d_o} g_o$$

we see that $\deg_1(S) < d$

Proof of Extension Theorem

- Now we are ready to prove that $I_a = (g_o(x_1, \mathbf{a}))$ by induction.
- By the above, claim is true for any $g_j \in G$ with $d_j < d_o$.
- Let $d \geq d_o$ and assume claim is true for any $g_j \in G$ with $d_j < d$.
- Let $g_i \in G$ be such that $d_i = d$.
- Taking standard representation of $S(g_i, g_o) = \sum_{k=1}^{t} B_k g_k$, where

$$S := c_o g_j - c_j x_1^{d-d_o} g_o$$

  we see that $\deg_1(S) < d$
- Thus, if $B_k c_k \neq 0$ then $\deg_1(g_k(x_1, x)) < d$, which by induction implies

$$g_k(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a})) \Rightarrow S \in (g_o(x_1, \mathbf{a})) \Rightarrow g_j(x_1, \mathbf{a}) \in (g_o(x_1, \mathbf{a}))$$

  as $c_o(\mathbf{a}) \neq 0$.

Resultants - Another Proof of Extension Theorem

- Univariate question: given two polynomials $f, g \in \mathbb{F}[x]$, when will they have a common root?

Resultants - Another Proof of Extension Theorem

- Univariate question: given two polynomials $f, g \in \mathbb{F}[x]$, when will they have a common root?
- As $\mathbb{F}[x]$ is an *Euclidean domain*, we have:

$$\gcd(f(x), g(x)) = 1 \iff$$
$$\exists\ s(x), t(x) \in \mathbb{F}[x] \text{ s.t. } s(x) \cdot f(x) + t(x) \cdot g(x) = 1$$

$\hookrightarrow < \deg f + \deg g$

$\deg(s) < \deg(g)$

$\deg(t) < \deg(f)$

$\overbrace{(s - ga)}^{< \deg(g)}$ $\qquad \overbrace{(t + fa)}^{< \deg f}$ is also solution $a \in \mathbb{F}[x]$

Resultants - Another Proof of Extension Theorem

- Univariate question: given two polynomials $f, g \in \mathbb{F}[x]$, when will they have a common root?
- As $\mathbb{F}[x]$ is an *Euclidean domain*, we have:

$$\gcd(f(x), g(x)) = 1 \Leftrightarrow$$
$$\exists\ s(x), t(x) \in \mathbb{F}[x] \text{ s.t. } s(x) \cdot f(x) + t(x) \cdot g(x) = 1$$

- We can also assume w.l.o.g. that $\deg(s) < \deg(g)$ and $\deg(t) < \deg(f)$.
- Viewing the equation $s(x) \cdot f(x) + t(x) \cdot g(x) = 1$ as a linear system, we have:

$$\begin{cases} s_0 \cdot f_0 + t_0 \cdot g_0 = 1 & \text{constant coefficient} \\ \sum_{i=0}^{k} s_i \cdot f_{k-i} + t_i \cdot g_{k-i} = 0 & \text{coefficient of degree } k \end{cases}$$

$f_i, g_i \in \mathbb{F}$

Sylvester Matrix & Resultant

- In matrix form (for simplicity $\deg(f) = 3, \deg(g) = 2$):

$$\begin{pmatrix} f_0 & 0 & g_0 & 0 & 0 \\ f_1 & f_0 & g_1 & g_0 & 0 \\ f_2 & f_1 & g_2 & g_1 & g_0 \\ f_3 & f_2 & 0 & g_2 & g_1 \\ 0 & f_3 & 0 & 0 & g_2 \end{pmatrix} \cdot \begin{pmatrix} s_0 \\ s_1 \\ t_0 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

s, t

Sylvester matrix

$$f(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3$$

$$g(x) = g_0 + g_1 x + g_2 x^2$$

Sylvester Matrix & Resultant

- In matrix form (for simplicity $\deg(f) = 3, \deg(g) = 2$):

$$\begin{pmatrix} f_0 & 0 & g_0 & 0 & 0 \\ f_1 & f_0 & g_1 & g_0 & 0 \\ f_2 & f_1 & g_2 & g_1 & g_0 \\ f_3 & f_2 & 0 & g_2 & g_1 \\ 0 & f_3 & 0 & 0 & g_2 \end{pmatrix} \cdot \begin{pmatrix} s_0 \\ s_1 \\ t_0 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Definition (Sylvester Matrix)

The matrix arising from the linear system is called *Sylvester Matrix*. It is denoted by

$$Syl_x(f, g)$$

Sylvester Matrix & Resultant

- In matrix form (for simplicity $\deg(f) = 3, \deg(g) = 2$):

$$\begin{pmatrix} f_0 & 0 & g_0 & 0 & 0 \\ f_1 & f_0 & g_1 & g_0 & 0 \\ f_2 & f_1 & g_2 & g_1 & g_0 \\ f_3 & f_2 & 0 & g_2 & g_1 \\ 0 & f_3 & 0 & 0 & g_2 \end{pmatrix} \cdot \begin{pmatrix} s_0 \\ s_1 \\ t_0 \\ t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Definition (Sylvester Matrix)

The matrix arising from the linear system is called *Sylvester Matrix*. It is denoted by

$$Syl_x(f, g)$$

Definition (Resultant)

The *Resultant* of f, g is the determinant of the Sylvester Matrix:

$$\mathrm{Res}_x(f, g) = \det(Syl_x(f, g))$$

Sylvester Matrix - General Case

$$f(x) = f_0 + f_1 x + \cdots + f_m x^m$$

$$g(x) = g_0 + g_1 x + \cdots + g_n x^n$$

$$\mathrm{Syl}_x(f,g) = \underbrace{\begin{pmatrix} f_0 & & \\ f_1 & \ddots & f_0 \\ \vdots & \ddots & \vdots \\ f_m & & \vdots \\ & \ddots & \\ & & f_m \end{pmatrix}}_{n} \underbrace{\begin{pmatrix} g_0 & \cdots & \\ g_1 & & g_0 \\ \vdots & & \vdots \\ g_n & & \vdots \\ & \ddots & \\ & & g_n \end{pmatrix}}_{m}$$

Resultants - Properties

- Resultant between two polynomials f, g is an *algebraic invariant*, and it is very important in computational algebra and algebraic geometry
- An important property is that the resultant is a *polynomial* over the *coefficients of* f, g

Resultants - Properties

- Resultant between two polynomials f, g is an *algebraic invariant*, and it is very important in computational algebra and algebraic geometry
- An important property is that the resultant is a *polynomial* over the *coefficients of* f, g
- From previous slides, another property is:

$$\mathrm{Res}_x(f, g) \neq 0 \iff \gcd(f, g) = 1 \quad \text{over } \mathbb{F}[x]$$

x

x^2

$\vdots$

x^{m+n-1}

Resultants - Properties

- Resultant between two polynomials f, g is an *algebraic invariant*, and it is very important in computational algebra and algebraic geometry
- An important property is that the resultant is a *polynomial* over the *coefficients of* f, g
- From previous slides, another property is:

$$\text{Res}_x(f,g) \neq 0 \;\Leftrightarrow\; \gcd(f,g) = 1 \quad \text{over } \mathbb{F}[x]$$

- Another important property is that, in some nice cases, the resultant behaves well under certain homomorphisms.

Let $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ be such that $\deg_1(f) = \ell$ and $\deg_1(g) = m$.
If $\mathbf{a} \in \mathbb{F}^{n-1}$ satisfies:

$\mathbb{F}[x_1, \bar{x}] \to \mathbb{F}[x_1]$

1. $\deg(f(x_1, \mathbf{a})) = \ell$
2. $g(x_1, \mathbf{a})$ is non-zero of degree $p \leq m$

and if $c(x_2, \ldots, x_n)$ is the leading coefficient of f, we have:

$\in \mathbb{F}[x_1]$

$$\text{Res}_{x_1}(f,g)(\mathbf{a}) = c(\mathbf{a})^{m-p} \cdot \text{Res}_{x_1}(f(x_1,\mathbf{a}), g(x_1,\mathbf{a}))$$

$\in \mathbb{F}[x_2, \ldots, x_n]$ $\quad \neq 0$

Discriminant

- A particular case which you have seen before is the discriminant.

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$
- That is, the polynomials $f(x)$ and $f'(x)$ have a common root.

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$
- That is, the polynomials $f(x)$ and $f'(x)$ have a common root.
- This implies that $x - \alpha \mid \gcd(f, f')$

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$
- That is, the polynomials $f(x)$ and $f'(x)$ have a common root.
- This implies that $x - \alpha \mid \gcd(f, f')$
- By the properties of the resultant, we have

$$\mathrm{Res}_x(f, f') = 0$$

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$
- That is, the polynomials $f(x)$ and $f'(x)$ have a common root.
- This implies that $x - \alpha \mid \gcd(f, f')$
- By the properties of the resultant, we have

$$\mathrm{Res}_x(f, f') = 0$$

- The *discriminant* of $f(x) \in R[x]$ is given by

$$\mathrm{disc}_x(f) := \mathrm{Res}_x(f, f')$$

Discriminant

- A particular case which you have seen before is the discriminant.
- From calculus, we know that $f(x) \in \mathbb{R}[x]$ has a *double root* $\alpha \in \mathbb{R}$ iff α is a root of $f(x)$ and of $f'(x)$
- That is, the polynomials $f(x)$ and $f'(x)$ have a common root.
- This implies that $x - \alpha \mid \gcd(f, f')$
- By the properties of the resultant, we have

$$\mathrm{Res}_x(f, f') = 0$$

- The *discriminant* of $f(x) \in R[x]$ is given by

$$\mathrm{disc}_x(f) := \mathrm{Res}_x(f, f')$$

- Why is it called discriminant? If $f(x) = ax^2 + bx + c$, we get

$$\mathrm{disc}_x(f) = -a \cdot (b^2 - 4ac)$$

Does this look familiar? :)

Extension Theorem

- *Extension Theorem*

  Let $\mathbb{F}$ be an *algebraically closed* field, $I := (f_1, \ldots, f_s) \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and let I_1 be the first elimination ideal of I. For each $1 \leq i \leq s$, write f_i as

  $$f_i = c_i(x_2, \ldots, x_n) \cdot x_1^{d_i} + \text{ lower degree terms in } x_1$$

  where c_i's are non-zero and $d_i \geq 0$. If

  $$(a_2, \ldots, a_n) \in V(I_1)$$

  that is, it is a partial solution, and if

  $$(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$$

  then there is $a_1 \in \mathbb{F}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.

Extension Theorem

- *Extension Theorem*

  Let $\mathbb{F}$ be an *algebraically closed* field, $I := (f_1, \ldots, f_s) \subseteq \mathbb{F}[x_1, \ldots, x_n]$ and let I_1 be the first elimination ideal of I. For each $1 \leq i \leq s$, write f_i as

  $$f_i = c_i(x_2, \ldots, x_n) \cdot x_1^{d_i} + \text{ lower degree terms in } x_1$$

  where c_i's are non-zero and $d_i \geq 0$. If

  $$(a_2, \ldots, a_n) \in V(I_1)$$

  that is, it is a partial solution, and if

  $$(a_2, \ldots, a_n) \notin V(c_1, \ldots, c_s)$$

  then there is $a_1 \in \mathbb{F}$ such that $(a_1, a_2, \ldots, a_n) \in V(I)$.
- Extension step fails then the leading coefficients must vanish

Resultants and Extension Theorem

- Similarly to the previous proof we know that the ideal

$$I_{\mathbf{a}} := \{f(x_1, \mathbf{a}) \mid f \in I\} \subseteq \mathbb{F}[x_1]$$

is generated by some polynomial $g(x_1, \mathbf{a}) \in \mathbb{F}[x_1]$, where $g \in I$, as $\mathbb{F}[x_1]$ is PID.

Resultants and Extension Theorem

- Similarly to the previous proof we know that the ideal

$$I_{\mathbf{a}} := \{f(x_1, \mathbf{a}) \mid f \in I\} \subseteq \mathbb{F}[x_1]$$

  is generated by some polynomial $g(x_1, \mathbf{a}) \in \mathbb{F}[x_1]$, where $g \in I$, as $\mathbb{F}[x_1]$ is PID.

- $\mathbf{a} \notin V(c_1, \ldots, c_s)$ implies that for some $i \in [s]$, we have $c_i(\mathbf{a}) \neq 0$. Thus, we know that ~~$g(x_1)$~~ $g(x_1, \bar{a})$ is non-zero.

I_a not zero ideal

$f_i(x_1, \bar{a}) \neq 0 \in I_a$

Resultants and Extension Theorem

- Similarly to the previous proof we know that the ideal

$$I_{\mathbf{a}} := \{f(x_1, \mathbf{a}) \mid f \in I\} \subseteq \mathbb{F}[x_1]$$

  is generated by some polynomial $g(x_1, \mathbf{a}) \in \mathbb{F}[x_1]$, where $g \in I$, as $\mathbb{F}[x_1]$ is PID.
- $\mathbf{a} \notin V(c_1, \ldots, c_s)$ implies that for some $i \in [s]$, we have $c_i(\mathbf{a}) \neq 0$. Thus, we know that $g(x_1)$ is non-zero.

  $f = f_i(x_1, \bar{x})$
- Let $h(\mathbf{x}) = \mathrm{Res}_{x_1}(f, g) \in I_1$
- We know that $h(\mathbf{a}) = 0$, since $\mathbf{a} \in V(I_1)$

$$\mathrm{Res}_{x_1}(f, g) = s \cdot f + t \cdot g$$

$$\in I \cap \mathbb{F}[x_2, \ldots, x_n]$$

Resultants and Extension Theorem

- Similarly to the previous proof we know that the ideal

$$I_{\mathbf{a}} := \{ f(x_1, \mathbf{a}) \mid f \in I \} \subseteq \mathbb{F}[x_1]$$

  is generated by some polynomial $g(x_1, \mathbf{a}) \in \mathbb{F}[x_1]$, where $g \in I$, as $\mathbb{F}[x_1]$ is PID.
- $\mathbf{a} \notin V(c_1, \ldots, c_s)$ implies that for some $i \in [s]$, we have $c_i(\mathbf{a}) \neq 0$. Thus, we know that $g(x_1)$ is non-zero.
- Let $h(\mathbf{x}) = \mathrm{Res}_{x_1}(f, g) \in I_1$
- We know that $h(\mathbf{a}) = 0$, since $\mathbf{a} \in V(I_1)$
- By property of Resultant, and the fact that the degree of f did not drop, there is $a_1 \in \mathbb{F}$ such that $f(a_1, \mathbf{a}) = g(a_1, \mathbf{a}) = 0$

$$0 = \mathrm{Res}_{x_1}(f, g)(\bar{a}) = c_i(\bar{a})^{\cdots} \cdot \mathrm{Res}_{x_1}(f(x_1, \bar{a}), g(x_1, \bar{a}))$$

$c_i(\bar{a}) \neq 0$

Resultants and Extension Theorem

- Similarly to the previous proof we know that the ideal

$$I_{\mathbf{a}} := \{f(x_1, \mathbf{a}) \mid f \in I\} \subseteq \mathbb{F}[x_1]$$

  is generated by some polynomial $g(x_1, \mathbf{a}) \in \mathbb{F}[x_1]$, where $g \in I$, as $\mathbb{F}[x_1]$ is PID.
- $\mathbf{a} \notin V(c_1, \ldots, c_s)$ implies that for some $i \in [s]$, we have $c_i(\mathbf{a}) \neq 0$. Thus, we know that $g(x_1)$ is non-zero.
- Let $h(\mathbf{x}) = \text{Res}_{x_1}(f, g) \in I_1$
- We know that $h(\mathbf{a}) = 0$, since $\mathbf{a} \in V(I_1)$
- By property of Resultant, and the fact that the degree of f did not drop, there is $a_1 \in \mathbb{F}$ such that $f(a_1, \mathbf{a}) = g(a_1, \mathbf{a}) = 0$
- Since $I_a = (g(x_1, \mathbf{a}))$, if a_1 is a root of $g(x_1, \mathbf{a})$ then it is a root of any polynomial in I_a and thus $(a_1, \mathbf{a})$ is a solution.

Conclusion

- Today we learned about Elimination and Extension Theorems
- These results allow us to solve systems of polynomial equations
- Saw how Groebner bases (w.r.t. lex order) behave nicely with respect to elimination
- Saw how Groebner bases can help us extend partial solutions
- Learned about Resultant, and how it can also help us in the Extension Theorem

Acknowledgement

- Lecture based entirely on the book by CLO: Ideals, varieties and algorithms (see course webpage for a copy - or get online version through UW library)