Lecture 3
Algebraic Computation

Chee Yap
Courant Institute of Mathematical Sciences
New York University

Overview

We introduce some basic concepts of algebraic computation.

- 0. Review
- I. Algebraic Preliminaries
- II. Resultants and Algebraic Numbers
- III. Sturm Theory

0. REVIEW

ANSWERS and DISCUSSIONS

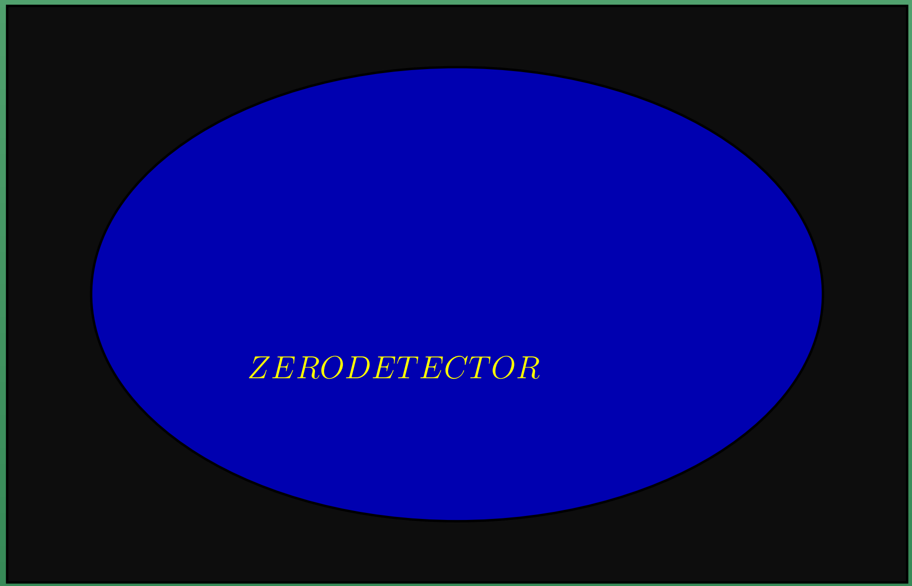
- Your experience with CORE so far?
- It did not print 11 digits of $\sqrt{2}$ because...
  - To fix it, you do ...
- Exercise on Implementation of Convex Hull
  - Send to Sung-il Pae (T.A.) your solutions, and he will reply with the answers.

What is EGC? Now you know...

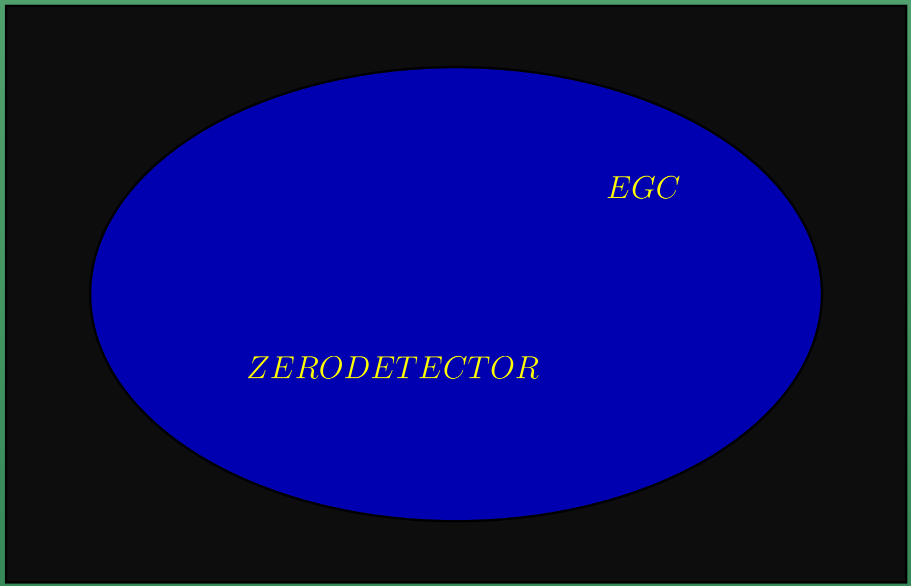
- Numerical Nonrobustness is widespread
- It has many negative impact on productivity and automation
- EGC prescribes that we compute the exact geometric relations to ensure consistency
  - ∗ Just take the right branch!
- It is the most successful approach
  - ∗ Can duplicate results of any other approach!
- EGC principles can be achieved by using a general library like CORE

- EGC can be expensive, but an effective technique is the use of filters and generalization
  - ∗ For bounded-depth rational problems, this is a small constant factor
  - ∗ E.g., convex hulls, line arrangements, etc, in low dimensions
- The center piece of any EGC libraries is an approximate evaluation algorithm for expressions
- The center of this algorithm is a Zero Detector

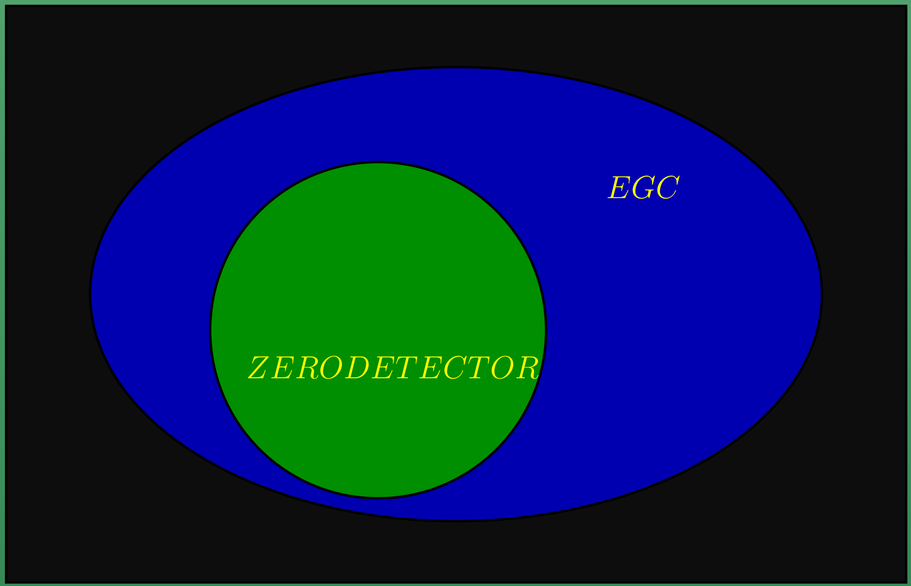
- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...
- EXERCISE



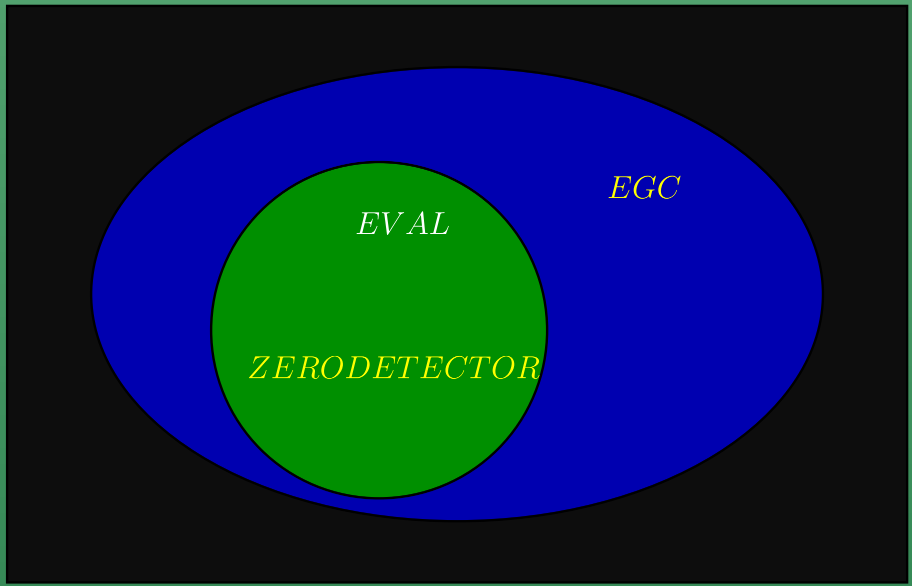
- Many challenges of EGC remain:
    - ∗ efficiency issues (zero bounds, filters and beyond)
    - ∗ geometric rounding
    - ∗ theory of EGC
    - ∗ transcendental computation, ...
- EXERCISE



- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...
- EXERCISE

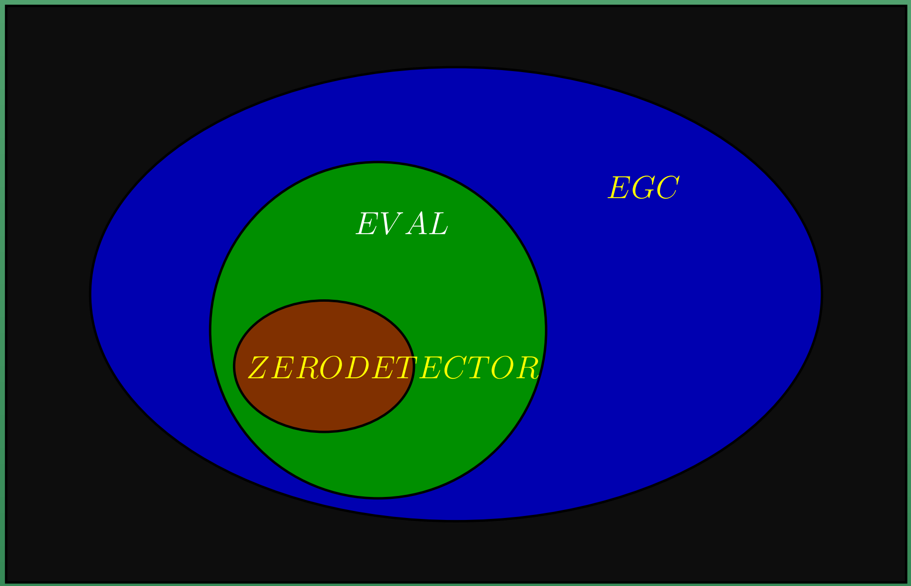


- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...
- EXERCISE

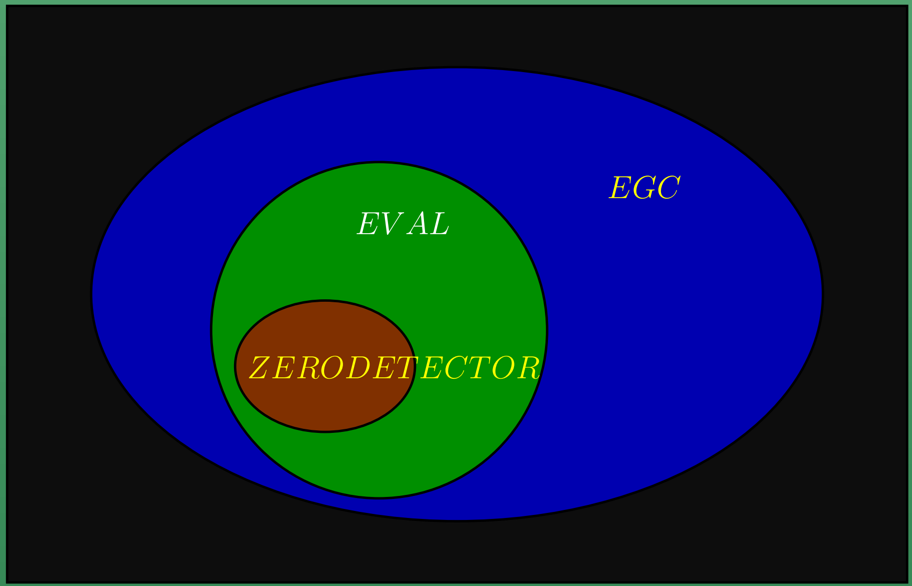


- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...

- EXERCISE



- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...
- EXERCISE



- Many challenges of EGC remain:
  - ∗ efficiency issues (zero bounds, filters and beyond)
  - ∗ geometric rounding
  - ∗ theory of EGC
  - ∗ transcendental computation, ...
- EXERCISE

$*$ Let the point p be given as the intersection of two lines, $p = L \cap L'$ where L, L' are given by their equations. If we want to compute $\widetilde{p}$ to s-bits of relative precision, what is the precision necessary in the coefficients of L and L'?

Algebraic Preliminaries

- What is between $\mathbb{Q}$ and $\mathbb{R}$?

- $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{A} \subseteq \mathbb{R} \subseteq \mathbb{C}$

  * Ring has $+, -, \times$ and $0, 1$. E.g., $\mathbb{Z}$
  * Field is a ring with $\div$. E.g., $\mathbb{Q}$
  * Domain: a ring where $xy = 0$ implies $x = 0$ or $y = 0$ (no zero divisor)
  * Ring is commutative if $xy = yx$. Assume this unless otherwise noted!

- Some Constructions in Algebra

  * Field $F \subseteq$ Domain $D \subseteq$ Ring R
  * Ring $R \subseteq R[X] \subseteq R[X, Y] \subseteq \ldots$

- ∗ Domain $D \subseteq$ Quotient field $Q_D \subseteq$ Algebraic closure $\overline{D}$

- ∗ Special case: $R[X] \Rightarrow R(X)$
- ∗ Ring R to matrix ring $R^{n\times n}$

- Polynomial $A(X) \in R[X]$ of degree m:
  - ∗ $A(X) = \sum_{i=0}^{m} a_i X^i$, $(a_m \neq 0)$
  - ∗ Leading coefficient, $a_m \neq 0$
  - ∗ $A(X)$ is monic if $a_m = 1$
  - ∗ Zero or root of $A(X)$: any $\alpha \in R$ such that $A(\alpha) = 0$

- Size measures for $A(X) \in \mathbb{C}[X]$
  - ∗ $\|A\|_k := \sqrt[k]{\sum_{i=0}^{m} |a_i|^k}$
  - ∗ Height of A is $\|A\|_\infty$

* Length of A is $\|A\|_2$

- Fundamental Theorem of Algebra:

    * A polynomial $A(X) \in \mathbb{C}[X]$ of degree m has exactly m zeros

    * i.e., $A(X) = a_m \prod_{i=1}^{m}(X - \alpha_i)$

- UFD: Unique factorization domain

    * $u \in D$ is a unit if if u has an inverse

    * Two elements $a, b \in D$ are associates if $a = ub$ for some unit u

    * a is irreducible if the only element that divides a is a unit or an associate of a

    * D is UFD if all non-zero $a \in D$ is equal to a product of

irreducibles, up to associates

- Fundamental Theorem of Arithmetic: $\mathbb{Z}$ is a UFD
  - * GAUSS LEMMA: if D is a UFD then so is $D[X]$

NOTE: A field is always a UFD

- GCD: Greatest Common Divisor
  - * In a UFD, we can define $\mathtt{GCD}(a, b)$
  - * We compute GCD's in $\mathbb{Z}$ and in $\mathbb{Q}[X]$ by Euclid's algorithm
  - * GCD over $\mathbb{Z}[X]$ is slightly trickier

- QUESTIONS
  - * From the above examples, show a ring that is not a domain.

* From the above examples, show a non-commutative ring.

* Prove that $\sqrt{x} + \sqrt{y}$ is an algebraic integer if x, y are positive integers

* What are the units in a field?

Algebraic Numbers

- The zero α of an integer polynomial $A(X) \in \mathbb{Z}[X]$ is called an algebraic number
  - $*$ If $A(X)$ is monic, α is an algebraic integer
  - $*$ NOTE: If $\alpha \in \mathbb{Q}$ is an algebraic integer, then $\alpha \in \mathbb{Z}$

- Let $A(X) \in \mathbb{Z}[X]$
  - $*$ $A(X)$ is primitive if the coefficients of $A(X)$ have no common factor except ± 1
  - $*$ Can always write $A(X) = c \cdot B(X)$ where $c \in \mathbb{Z}$ and $B(X) \in \mathbb{Z}[X]$ is primitive

- The minimal polynomial of α is the primitive polynomial in $\mathbb{Z}[X]$ of minimal degree.
  - $*$ It is basically unique

* Degree and height of α is the degree and height of this minimal polynomial

Resultants

- Resultants is a very important constructive tool for manipulation of algebraic numbers

- Let D be any UFD (e.g., $D = \mathbb{Z}$ or $D = \mathbb{Q}[X]$)

- Let $A(X) \in \sum_{i=0}^{m} a_i X^i, B(X) \in \sum_{j=0}^{n} b_i X^i$ be polynomials in $D[X]$, $a_m b_n \neq 0$

- The resultant $\texttt{res}(A, B)$ of A, B is the determinant of the Sylvester matrix of A, B:
  - This is a $(m+n) \times (m+n)$ matrix $Syl(A, B)$

[17]

$$Syl(A,B) = \begin{bmatrix} a_m & a_{m-1} & \cdots & & a_0 & & & \\ & a_m & a_{m-1} & \cdots & & a_0 & & \\ & & \ddots & & & & \ddots & \\ & & & a_m & a_{m-1} & \cdots & & a_0 \\ b_n & b_{n-1} & \cdots & b_1 & b_0 & & & \\ & b_n & b_{n-1} & \cdots & b_1 & b_0 & & \\ & & \ddots & & & & \ddots & \\ & & & b_n & b_{n-1} & \cdots & & b_0 \end{bmatrix}$$

- LEMMA: $\mathtt{GCD}(A,B) \notin D$ iff $\mathtt{res}(A,B) = 0$
  * Sketch: Set up "$\mathtt{GCD}(A,B) \notin D$" as a system of equations involving $Syl(A,B)$

- Now assume $D = \mathbb{C}$
  * So $A(X) = a\prod_{i=1}^{m}(X-\alpha_i)$ and $B(X) = b\prod_{j=1}^{n}(X-\beta_j)$

- THEOREM A: The resultant $\texttt{res}(A, B)$ is equal to each of the following
  - (A) $a^n \prod_{i=1}^{m} B(\alpha_i)$
  - (B) $(-1)^{mn} b^m \prod_{j=1}^{n} A(\beta_j)$
  - (C) $a^n b^m \prod_{i=1}^{m} \prod_{j=1}^{n} (\alpha_i - \beta_j)$

- COROLLARY:
  - (D) $\beta_j \pm \alpha_i$ is a zero of $D(X) = \texttt{res}_Y(A(Y), B(X \mp Y))$
  - (E) $\alpha_i \beta_j$ is a zero of $E(X) = \texttt{res}_Y(A(Y), Y^n B(X/Y))$
  - (F) $1/\alpha_i$ is a zero of $F(X) = X^m A(1/X)$

- COROLLARY:
  - The algebraic integers form a ring
  - The algebraic numbers form a field

- THEOREM: If $\alpha_0, \ldots, \alpha_m$ are algebraic numbers, then any root of $\sum_{i=0}^{m} \alpha_i X^i$ is also algebraic

* The proof uses theory of symmetric functions

Zero Bounds and Separation Bounds

- Cauchy Bound: Suppose α is the zero of $A(X) = \sum_{i=0}^{m} a_i X^i \in \mathbb{Z}[X]$
  - * Then $|\alpha| \leq (1 + H)$ where $H = \|A\|_\infty$
- Pf: If $|\alpha| \leq 1$, the result is true. Assume otherwise.
  - * Then $|a_m| \cdot |\alpha|^m \leq H \sum_{i=0}^{m-1} |\alpha^i| = H(|\alpha|^m - 1)/(|\alpha| - 1) < H|\alpha|^m/(|\alpha| - 1)$.
  - * The claim follows. QED
- Corollary: $|\alpha| \geq 1/(1 + H)$
  - * Pf: Note that $1/|\alpha|$ is the zero of $B(X) = X^m A(1/X)$.
  - * But the height of $B(X)$ is also H. QED
- Constructive Zero Bounds
  - * Based on the structure of the expression (see Exercise)

- Root Separation Bounds
  - $*$ Define $\texttt{Sep}(A)$ to be the minimum of $|\alpha - \beta|$ where α, β range over all pairs of distinct zeros of $A(X)$

- Discriminant of $A(X)$ is defined as $a^{-1}\texttt{res}(A, A')$ where a is A's leading coefficient
  - $*$ Check: If $A(X) \in D[X]$ then $\texttt{Disc}(A) \in D[X]$

- THEOREM: Let $\alpha_1, \ldots, \alpha_m$ are all the complex roots of $A \in \mathbb{C}[X]$, not necessarily distinct. Up to sign, the following three quantities are equal:
  - $*$ (A) $a^{-1}\texttt{res}(A, A')$ where a is A's leading coefficient
  - $*$ (B) $\prod_{1 \le i < j \le m} (\alpha_i - \alpha_j)^2$
  - $*$ (C) the square of the determinant of the Vandermonde

matrix,

$$V_m(\alpha_1, \alpha_2, \ldots, \alpha_m) := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_m^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{m-1} & \alpha_2^{m-1} & \cdots & \alpha_m^{m-1} \end{bmatrix}$$

- THEOREM (Mahler)
  - Then $\texttt{Sep}(A) > \sqrt{|disc(A)|} \cdot m^{-(m/2)+1} M(A)^{1-m}$

where $M(A)$ is Mahler measure.

PROOF: Result is trivial when A has multiple roots, for then $\texttt{Disc}(A) = 0$. Else,

assume $\texttt{Sep}(A) = |\alpha_1 - \alpha_2|$ where $|\alpha_1| \geq |\alpha_2|$.

Starting with the Vandermonde matrix, we may subtract the second column from the first column, preserving the determinant.

The first column (transposed) is now $(0, \alpha_1 - \alpha_2, \alpha_1^2 - \alpha_2^2, \ldots, \alpha_1^{m-1} - \alpha_2^{m-1}) = (\alpha_1 - \alpha_2)(0, 1, \alpha_1 + \alpha_2, \ldots, \sum_{i=0}^{m-2} \alpha_1^i \alpha_2^{m-2-i})$.

The 2-norm of $(0, 1, \alpha_1 + \alpha_2, \ldots, \sum_{i=0}^{m-2} \alpha_1^i \alpha_2^{m-2-i})$ is at most $\sqrt{\sum_{i=0}^{m-2}(i+1)^2 |\alpha_1|^i}$.

Hence this 2-norm is at most $h_1 := \sqrt{m^3/3}\max\{1, |\alpha_1|\}^{m-1}$.

By Hadamard's bound, the Vandermonde determinant is at most $\mathtt{Sep}(A)\prod_{i=1}^m h_i$ where h_i is any upper bound on 2-norm of the ith column.

We have already computed h_1. For $i \geq 2$, we can choose $h_i = \sqrt{m}\max\{1, |\alpha_i|\}^{m-1}$.

The product of these bounds yields $\sqrt{|\mathtt{Disc}(A)|} < \mathtt{Sep}(A) m^{(m/2)+1} \prod_{i=1}^m |\max\{1, |\alpha_i|\}^{m-1} = \mathtt{Sep}(A) m^{(m/2)+1} M(A)$

The conclusion of the theorem is now clear.

- EXERCISE
  - Using Theorem A above, give height bounds for $\alpha\beta$ and $\alpha \pm \beta$, assuming we know heights and degree bounds for α, β

Sturm Theory

- Now assume $A, B \in \mathbb{R}[X]$ and $\deg A > \deg B > 0$
  - $*$ The generalized Sturm sequence for (A, B) is $(A_0, A_1, \ldots, A_h)$ where $(A_0, A_1) = (A, B)$ and $A_{i+1} = -(A_{i-1} \bmod A_i)$, with $A_{h+1} = 0$

- Let $\mathbf{a} = (a_0, \ldots, a_h)$ where $a_i \in \mathbb{R}$
  - $*$ Let $\mathtt{Var}(\mathbf{a})$ be the number of sign variations in $\mathbf{a}$
  - $*$ E.g., $\mathtt{Var}(1, 0, -1, 0, 3) = 2$ and $\mathtt{Var}(0, 8, 1, 0, 4, -3, 0) = 1$
  - $*$ Write $\mathtt{Var}_{A,B}(a)$ for $\mathtt{Var}(A_0(a), A_1(a), \ldots, A_h(a))$

- THEOREM (Sturm): If $B = A'$, then for all $a < b$ such that $A(a)A(b) \neq 0$
  - $*$ Then $\mathtt{Var}_{A,B}(a) - \mathtt{Var}_{A,B}(b)$ is equal to the number of

real roots of A in $[a, b]$.

PROOF: First assume (A, B) has no common zero.

Let $c \in [a, b]$ and $v_i(c) := \mathtt{Var}(A_{i-1}(c), A_i(c), A_{i+1}(c))$ for $i = 0, \ldots, h$.

(a) $V_{i-1}(c) = V_i(c) = 0$ implies $V_{i-2}(c) = V_{i+1}(c) = 0$

(b) So $A_h(c) \neq 0$ (otherwise c is common zero of A, B)

(c) From (a), $V_{i-1}(c)^2 + V_{i+1}(c)^2 \neq 0$ for $1 < i < h$.

(d) This implies $2\mathtt{Var}_{A,B}(c) = \sum_{i=0}^{h} v_i(c)$

(e) If $i > 0$ and $A_i(c) = 0$ then $v_i(c^-) = v_i(c^+)$.

(f) Hence $v_i(c)$, and so $\mathtt{Var}_{A,B}(c)$ does not change when c passes through a zero of A_i $(i > 0)$

(g) If $A_0(c)$ then $v_0(c)$ decreases by 1 (use the fact that $B = A'$)

(h) Thus, $\mathtt{Var}_{A,B}(c)$ decreases by 1 each time as c passes over a zero of A, but does not change otherwise.

(i) This implies $\mathtt{Var}_{A,B}(a) - Val_{A,B}(c)$ equals the number

of real zeros of A in $[a, b]$.
Finally, suppose $C = \mathtt{GCD}(A, B)$ has degree > 0. The sequence $(A_0/C, A_1/C, \ldots, A_h/C)$ has the same properties as what we proved in (i).

- We can now isolate all the real zeros of a polynomial $A(X)$ using an obvious bisection
  * NOTe: All real zeros lies in the interval $[-1-H, 1+H]$ where H is the height of $A(X)$ Can extend Sturm sequence to find all complex roots (See Chapter 7 [Yap-Fundamental])

Conclusions

- Arithmetic on algebraic numbers are possible via resultant methods, but such methods are inefficient

- Algebraic numbers can be manipulated numerically and compared exactly if you know root bounds

EXERCISES

- Isolating Interval Representation (IIR):
  * A real algebraic number α can be represented by a pair $(A(X), [a, b])$ such that α is the only zero of $A(X) \in \mathbb{Z}[X]$ in $[a, b]$
- Show how to perform the four arithmetic operations on IIR's
- Show how to do comparisons on IIR's
- Compare the efficiency of IIR's to our expression approach

REFERENCE

- Chapter 6 of [Yap-FundamentalProblems], on roots of polynomials.

"A rapacious monster lurks within every computer, and it dines exclusively on accurate digits."
– B.D. McCullough (2000)

THE END