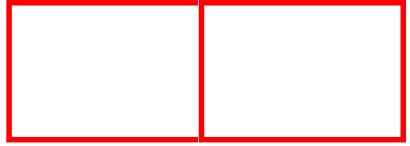


Computer Algebra

Basic Information

- Lecturer: Kyriakos Kalorkoti (KK).
- Email: `kk@inf.ed.ac.uk`
- Web: `http://www.inf.ed.ac.uk/teaching/courses/ca`
  - Lecture log here.

Working definition of Computer Algebra: Algorithms, techniques and tools to assist with mathematical work (not just numerical).

Syllabus

1. Introduction to Axiom (Exercises 1 and home reading, see lecture log for details).
2. Basic Structures and Algorithms.
3. Keeping the Data Small: Modular Methods.
4. Polynomial Simplification.
5. Real Roots of Polynomials.

Computer Algebra

Coursework

Accounts for 20% of final assessment.

Each assignment is allocated 3 weeks, they can all be done within 2 weeks at most.

1. Exploring Axiom (20%).
   - A timetabled way for you to get to know the system.
   - Write some simple code.
2. Computing with algebraic extensions (40%).
   - Key practical showing connection of abstract ideas with practical concerns.
   - Write some code.
   - Some pencil and paper parts also.
3. Operations on ideals (40%).
   - Uses ideas from the course with Axiom facilities as tools.
   - Do some calculations with Axiom (on ideals).
   - Some pencil and paper parts.
   - Do a past exam.

Coursework Submission

- Hard copy via ITO, code via `submit` command on DICE.
- See course web page for details.

General Study

- Allow around 2 hours of study per lecture.
- Try selected exercises from the notes: some will be suggested at the end of various lectures (and discussed at the next).
- Speak to me (or send email) if you need help; do this sooner rather than later.

Exam

- Recent past exams are a reasonable guide (but note change from Maple to Axiom).
- Revision will be much easier if you study continuously.
- If there is demand a meeting will be arranged to discuss approach to exam.
- A guide to revision will be issued at the end of the course.

My Educational Approach

1. The lecturer's job is to provide opportunities for students to learn the subject (lectures, notes, exercises, feedback, help).
2. The students' job is to use those opportunities to the full.
3. Attending lectures is *essential* not an optional extra.
4. Questions during lectures are strongly encouraged.

General Introduction: motivation

Eugene Wigner: The Unreasonable Effectiveness of Mathematics in the Natural Sciences.

> *The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure, even though perhaps also to our bafflement, to wide branches of learning.*

in Communications in Pure and Applied Mathematics, vol. 13, No. I (February 1960)

Obvious consequence: Developing tools to help is very important.

Observation: The usefulness of a deep concept is often far from obvious. The obvious approach is often not the best.

Examples

Differentiate:

$$f = \frac{32x^8 - 16x^7 + 82x^6 - 40x^5 + 85x^4 - 40x^3 + 101x^2 - 48x + 6}{8x^5 - 2x^4 + 4x^3 - x^2 + 12x - 3}.$$

Rule:

$$\frac{d(p/q)}{dx} = \frac{q\frac{dp}{dx} - p\frac{dq}{dx}}{q^2}.$$

Direct application gives a big mess!

- Use a machine.
- Simplify; in fact

$$f = 4x^3 - x^2 + 8x - 2.$$

Integrate

$$g = \frac{x^2 - 5}{x(x-1)^4}.$$

Decompose into partial fractions:

$$g = \frac{-5}{x} + \frac{5}{x-1} - \frac{5}{(x-1)^2} + \frac{6}{(x-1)^3} - \frac{4}{(x-1)^4}.$$

More generally consider

$$\frac{x+a}{x(x-b)(x^2+c).}$$

More ambitiously allow log, sin, cos, roots etc.

Find a formula for

$$\sum_{i=0}^{n} f(i)$$

where

- n is a symbol not a number.
- f comes from a fairly wide class of functions.

Examples:

$$\sum_{i=1}^{n} 2i^5 - i^3 + 1 = \frac{4n^6 + 12n^5 + 7n^4 - 6n^3 - 5n^2 + 12}{12}$$

$$\sum_{i=1}^{n} \frac{i+1}{2^i} = \frac{4 \cdot 2^n - n - 3}{2^n}.$$

Features of Computer Algebra Systems

- Interactive use.
- File handling.
- Polynomial manipulation.
- Elementary special functions.
- Arithmetic.
- Differentiation.
- Integration.
- Own programming language.
- Huge number of built in (mathematical etc.) data structures and algorithms.
- Graphics.

Example Axiom code

```
-- Returns the Cauchy bound on the positive roots of a polynomial
-- Inputs:
--   p.....a univariate polynomial with rational coefficients.
--   x.....a symbol.
-- Output:
--   An upper bound on the positive roots of the polynomial (an error is returned if
--   it obviously has none, but an absense of an error does not mean it has any).
-- Remark:
--   This is a quick and dirty version, it uses floating point arithmetic!
cauchy(p:POLY(FRAC(INT))):Float==
  local V,x,n
  V:=variables(p)
  if #(V)>1 then error "The input must be a univariate polynomial"
  else if #(V)=1 then x:=V.1
  if leadingCoefficient(p)<0 then p:=-p
  C:=coefficients(p)
  n:=0
  for c in C repeat if c<0 then n:=n+1
  if n=0 then error "The polynomial has no positive roots"
  else
    m:=degree(p,V).1
    lc:=leadingCoefficient(p)
    Blist:=[]::List(AlgebraicNumber)
    for i in 0..m-1 repeat
      coeff:=coefficient(p,x,i)
      if coeff<0 then Blist:=cons((-n*coeff/lc)^(1/(m-i)),Blist)
    B:=map(r+->r::Float,Blist)
    mx:=B.1::Float
    for v in B repeat if v::Float>mx then mx:=v::Float
    mx
```

Forward look to Exercise II

Problem: Find all Intersections of two Algebraic Curves
Requirement: Absolute reliability, no approximations.

Example:

- $f = x^5 + y^5 + 2\,y^3 - 1$.
- $g = x^2y^4 - xy^3 - 2$.

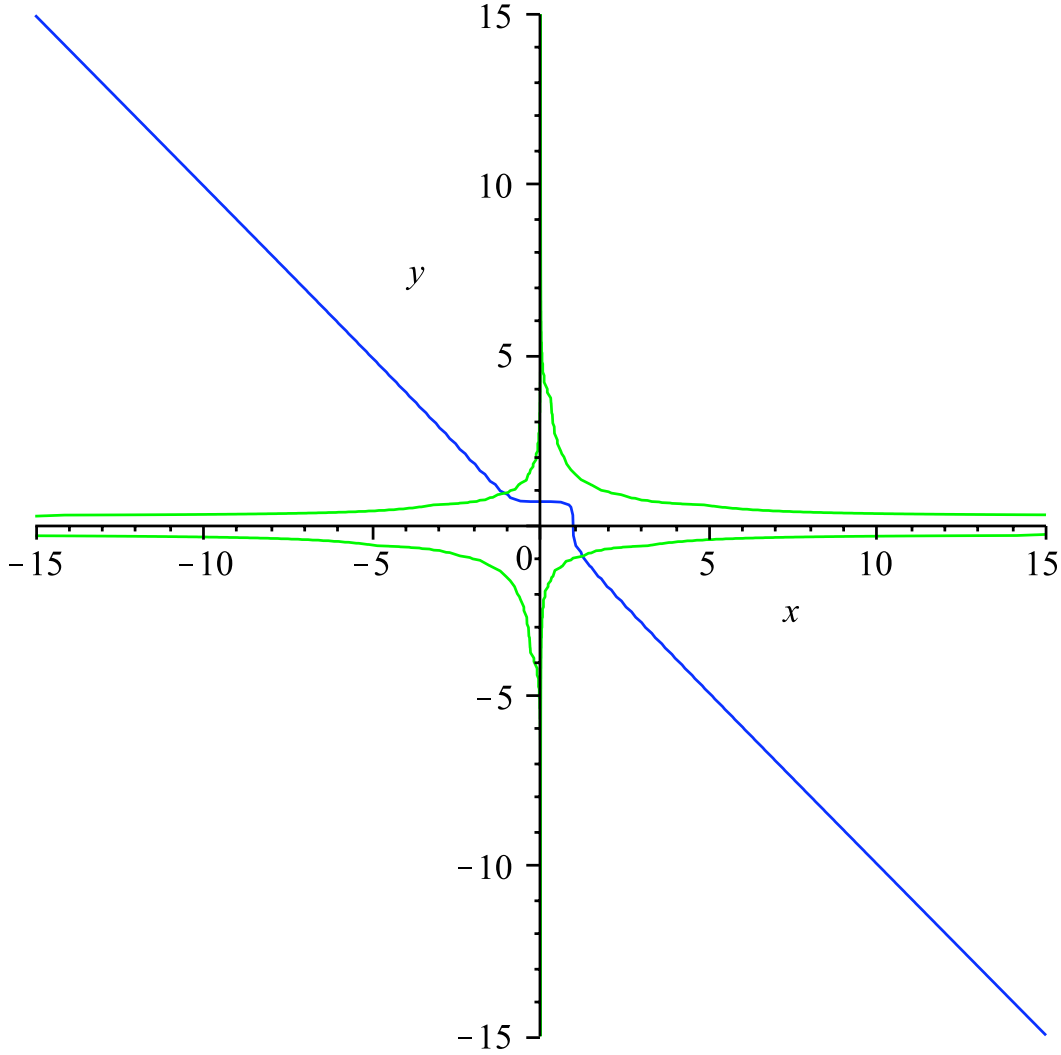
Fact: Common points are

$$(a, h(a))$$

where h is a polynomial of degree 29 and a is any solution of

$$a^{30} - 4\,a^{25} + 12\,a^{22} + 7\,a^{20} - 36\,a^{17} - 56\,a^{16} - 7\,a^{15} + 8\,a^{14} +$$
$$36\,a^{12} + 112\,a^{11} + 100\,a^{10} - 16\,a^9 - 64\,a^8 - 12\,a^7 - 56\,a^6 -$$
$$97\,a^5 - 120\,a^4 + 64\,a^3 + 64\,a^2 - 32 = 0.$$

Conclusion: The two curves have *exactly* 30 distinct common points (in the complex plane). Exactly 2 of them are real: approximately 1.254 and -1.098.


15
10
y
5
−15
−10
−5
0
5
10
15
x
−5
−10
−15

Algorithm: Short; discussed in Exercises 2.

Necessary tools: Some standard algebraic structures:

- polynomial rings,
- quotients of rings (special case),
- algebraic extensions and facts about them.

Implementation: Very straightforward in Axiom, quite short.

Basic Structures and Algorithms

Rationale

- Structures give us a convenient and precise language.
- They capture and abstract common patterns and properties.
- They have come about as the result of centuries of research by very many people.
- Exercises 2 will show convincingly how they help.

Standard Notation

1. $\mathbb{Z}$, the integers,
2. $\mathbb{Q}$, the rationals,
3. $\mathbb{R}$, the reals,
4. $\mathbb{C}$, the complex numbers,
5. $\mathbb{Z}_n$ the integers modulo n where $n \geq 1$ is a natural number.

Binary Operations

- Function on a set R taking two elements of R returning an element of R:
$$\circ : R \times R \to R.$$
- *Commutative* if
$$x \circ y = y \circ x \quad \text{for all } x, y \in R.$$
- *Associative* if
$$(x \circ y) \circ z = x \circ (y \circ z) \quad \text{for all } x, y, z \in R.$$

  Nice consequence:
$$x_1 \circ x_2 \circ \cdots \circ x_n$$
  gives same result for any valid bracketing. (*Prove* this.)

Rings

Ingredients

Set R with two binary operations $+$, $*$ called *addition* and *multiplication*.

Requirements

1. $+$ is associative,
2. $+$ is commutative,
3. there is an element 0 of R such that $x + 0 = x$ for all $x \in R$,
4. for each element x of R there is an element $y \in R$ such that $x + y = 0$, (i.e. x has an *additive inverse*),
5. $*$ is associative,
6. for all $x, y, z \in R$ we have

$$x * (y + z) = x * y + x * z, \quad (x + y) * z = x * z + y * z,$$

i.e., $*$ is left and right *distributive* over $+$.

Conventions & Facts

- $0 + x = x$ for all $x \in R$.
- 0 is unique.
- Additive inverse of an element x is *unique*, denoted by $-x$.
  - Write
    $$x - y$$
    instead of
    $$x + (-y).$$
- For all $x \in R$
  $$x * 0 = 0 = 0 * x.$$
- Usually write
  $$xy$$
  instead of
  $$x * y.$$

Examples of Rings

1. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ with usual addition & multiplication.
2. $2\mathbb{Z}$ set of all even integers, the usual addition & multiplication.
3. $\mathbb{Z}_n$ integers modulo integer $n \geq 1$. Addition & multiplication carried out as normal but take as result the remainder after division by n.
   More accurately elements are equivalence classes of remainders. Operations are on equivalence classes.
4. Square matrices of a fixed size with integer entries. Normal operations of matrix addition & matrix multiplication.
5. S any set, $P = \mathcal{P}(S)$ the *power set* of S (i.e., set of all subsets of S).
   - Addition is symmetric difference, i.e., $A + B$ is $A \cup B - A \cap B$.
   - Multiplication is intersection, i.e., $A * B$ is $A \cap B$.

   Example of a *Boolean ring*, i.e. $x * x = x$ for all x.

More Definitions

Ring R is:

- *commutative* if $*$ is commutative.
- has (multiplicative) *identity* if it has an element e s.t.

$$ex = xe = x, \quad \text{for all } x \in R.$$

Usually denote e by 1.

Note: Identity is unique if it exists.
If R has identity, say that x has (multiplicative) *inverse* if

$$xy = yx = 1, \quad \text{for some } y \in R.$$

Note: Inverse of an element x is unique if it exists, denoted by x^{-1}.

'Strange' behaviour: In $\mathbb{Z}_6$ we have $3 \neq 0$ and $2 \neq 0$ but

$$2 \times 3 = 0.$$

Not *really* strange, we are dealing with equivalence classes of remainders: 2×3 is divisible by 6, hardly a surprise!

Fields

Rings with extra properties:

1. there is a multiplicative identity that is *different* from 0,
2. multiplication is commutative,
3. every non-zero element has an inverse.

Here

$$xy = 0 \Rightarrow x = 0 \text{ or } y = 0.$$

Proof: Suppose $xy = 0$ but $x \neq 0$. Then x^{-1} exists. Thus

$$\begin{aligned} 0 &= x^{-1}0 \\ &= x^{-1}(xy) \\ &= (x^{-1}x)y \\ &= 1y \\ &= y \end{aligned}$$

'Strange' things still possible: $\mathbb{Z}_2$ is a field but

$$1 + 1 = 0.$$

Similar thing happens any *finite field*. Can also happen in infinite fields.

Examples of fields:

1. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ all with usual operations.
2. $\mathbb{Z}_p$ when p is a prime.

Note: $\mathbb{Z}_n$ is *not* a field if $n = 1$ or a composite number.

Suggested Exercise: 4.2.

Intermediate Structures

- **Integral domain:** (*ID*) commutative ring with identity, different from 0, s.t.

$$xy = 0 \Rightarrow x = 0 \text{ or } y = 0.$$

  Note: Every field is an ID (but not conversely).

  Consequence: If $ax = ay$ and $a \neq 0$ then $x = y$.

- **Unique factorization domain:** (*UFD*) notion of irreducible elements (cf prime numbers) and unique decomposition of elements like integer case.
  In UFD's *greatest common divisors* are guaranteed to exist.

Definition: Let $a, b \in R$, where R is a ring. We say that *a divides b*, written as $a \mid b$, if and only if $b = ac$ for some $c \in R$.

Note: Division is of no interest in fields.

Definition: Let $a, b \in D$, where D is an integral domain. Then

- d is a *common divisor* of a, b if $d \mid a$ and $d \mid b$.
- d is a *greatest common divisor (gcd)* of a, b if
  1. d is a common divisor of a, b and
  2. for all common divisors c of a, b we have $c \mid d$.

Note: If $a \neq 0$ or $b \neq 0$ then necessarily $d \neq 0$.
Also 0 is the only gcd of 0, 0.

Fact: If d_1, d_2 are two gcd's of a, b then there is an invertible element u of R s.t. $d_1 = ud_2$.
Conversely if d is a gcd of a, b and u is an invertible element of R then ud is also a gcd of a, b.

Canonical and Normal Representations

A representation is:

- *Canonical* if equality of objects is same as equality of representations.
  - Each object has exactly one representation.
- *Normal* if 0 has only one representation. (In a system with a notion of 0 and subtraction.)
  - This means that we can test objects for equality.

$$a = b \iff a - b = 0 \iff R(a - b) \equiv R(0).$$

Integers and Rationals

Integers

Use a large base B which

1. fits into a word (usually leave a bit for carries),
2. is usually a power of 2 or 10 and is largest power (of 2 or 10) s.t. B^2 representable in host machine arithmetic.

Representation: hold digits in a linked list or an array.

Karatsuba's Algorithm

Two integers of length n in base B:

$$x = aB^{n/2} + b,$$

$$y = cB^{n/2} + d,$$

(adjust appropriately for n odd). Now

$$xy = acB^n + (bc + ad)B^{n/2} + bd.$$

No improvement. *But*

$$bc + ad = (a + b)(c + d) - ac - bd.$$

Leads to time

$$t(n) = \begin{cases} k_1, & \text{if } n = 1; \\ 3t(n/2) + k_2 n, & \text{if } n > 1. \end{cases}$$

(k_1, k_2 constants). Solution:

$$t(n) = \Theta(n^{\log_2 3}), \quad (\log_2 3 \approx 1.67).$$

Pays off for integers of sufficiently many digits.

Fractions

Definition: a, b integers not both 0. *Greatest common divisor*, $\gcd(a, b)$, is largest integer d dividing both a and b.

Always represent a/b as p/q with $q \geq 1$ and $\gcd(p, q) = 1$. So can convert to an integer type if and only if $q = 1$.

Gives canonical form.

Representation: any structure that can hold a pair of integers.

Rational arithmetic

a/b, c/d in canonical form.

$$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd} = \frac{ac/\gcd(ac, bd)}{bd/\gcd(ac, bd)}$$

Much better:

$$d_1 = \gcd(a, d),$$
$$d_2 = \gcd(b, c).$$

Required canonical form is:

$$\frac{(a/d_1)(c/d_2)}{(b/d_2)(d/d_1)}.$$

Justified because

$$\gcd(a, b) = \gcd(c, d) = 1 \implies$$
$$\gcd(ac, bd) = \gcd(a, d)\gcd(b, c).$$

Division same.

For addition/subtraction put:

$$\frac{a}{b} \pm \frac{c}{d} = \frac{p}{q},$$

r.h.s. in canonical form.
Compute

$$p' = a\frac{d}{\gcd(b,d)} + c\frac{b}{\gcd(b,d)}$$
$$q' = \frac{bd}{\gcd(b,d)}$$

Now:

$$p = p'/\gcd(p',q'), \qquad q = q'/\gcd(p',q').$$

Suggested Exercise: 4.5

Euclid's Algorithm for the Integers

Simple properties of gcd's:

1. $\gcd(a, b) = \gcd(b, a)$.
2. $\gcd(a, b) = \gcd(|a|, |b|)$.
3. $\gcd(0, b) = |b|$.
4. $\gcd(a, b) = \gcd(a - b, b)$.

Simple (inefficient) algorithm ($a, b \geq 0$):

if $a = 0$ **then** b
elif $a < b$ **then** $\gcd(b, a)$
else $\gcd(a - b, b)$
fi

Improved version (Euclid's Algorithm)

Assume $a \geq 0$, $b > 0$ and put

$$a = qb + r, \quad 0 \leq r < b, \ q \in \mathbb{Z}.$$

q is *quotient* of a, b and r *remainder*. Have

$$\gcd(a, b) = \gcd(b, r).$$

Algorithm: Put $r_0 = a$, $r_1 = b$:

$$\begin{aligned} r_0 &= q_1 r_1 + r_2 \\ r_1 &= q_2 r_2 + r_3 \\ r_2 &= q_3 r_3 + r_4 \\ &\vdots \\ r_{s-2} &= q_{s-1} r_{s-1} + r_s \\ r_{s-1} &= q_s r_s + r_{s+1} \end{aligned}$$

where

$$r_{s+1} = 0 \quad \text{and} \quad 0 \leq r_i < r_{i-1}, \quad \text{for } 1 \leq i \leq s+1.$$

Extended version

Rewrite last step as

$$r_s = r_{s-2} - q_{s-1} r_{s-1}.$$

Remainder r_{s-1} can be written as

$$r_{s-1} = r_{s-3} - q_{s-2} r_{s-2}$$

so

$$r_s = -q_{s-1} r_{s-3} + (1 + q_{s-1} q_{s-2}) r_{s-2}.$$

Process can be continued until

$$r_s = u r_0 + v r_1$$

where u, v are *integers*.

Conclusion: If $d = \gcd(a, b)$ then there are integers u, v s.t.

$$d = ua + vb.$$

Can compute u, v by 'forwards' Euclid's algorithm.

Suggested Exercise: 4.9

Lemma: $\mathbb{Z}_n$ is a field if and only if n is a prime.

Polynomials

- R a commutative ring with 1.
- x a brand new symbol—called an *indeterminate* over R.
- *Polynomials* in *indeterminate* x with *coefficients* from R:

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots$$

  where $a_i \in R$ and all all but finitely many are 0
- Could just as well write

$$(a_0, a_1, a_2, \ldots)$$

  but x very useful.
- a_i is *coefficient* of x^i.
- a_0 is *constant term*.
- Set of all such polynomials denoted by $R[x]$.

Convenient abbreviation:

$$\sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

Equality:

$$\sum_{i=0}^{\infty} a_i x^i = \sum_{i=0}^{\infty} b_i x^i$$

iff

$$a_0 = b_0, \ a_1 = b_1, \ a_2 = b_2, \ldots$$

Sensible convention: write

$$2 + 5x^3 - 3x^5$$

instead of

$$2 + 0x + 0x^2 + 5x^3 + 0x^4 - 3x^5 + 0x^6 + \cdots$$

Turning $R[x]$ into a Ring

- Define $+$, $*$ on polynomials in the usual way.
- Makes $R[x]$ into commutative ring with 1.

Further definitions: For $p \in R[x]$ define:

- *Degree*, $\deg(p)$; undefined for zero polynomial.
- *Leading coefficient*, $\mathrm{lc}(p)$; undefined for zero polynomial.
- Basic facts:

$$\begin{aligned} \deg(p \pm q) &\leq \max\big(\deg(p), \deg(q)\big), \\ \deg(pq) &\leq \deg(p) + \deg(q), \\ \deg(pq) &= \deg(p) + \deg(q), \quad \text{if } \mathrm{lc}(p)\,\mathrm{lc}(q) \neq 0, \end{aligned}$$

whenever both sides defined.

Polynomial Functions

Given

$$p = a_0 + a_1 x + \cdots + a_n x^n$$

Define corresponding *function*

$$\hat{p} : R \to R,$$
$$\hat{p}(\alpha) = a_0 + a_1 \alpha + \cdots + a_n \alpha^n.$$

Note: p, $\hat{p}$ *very* different objects.
Consider equality of polynomials v. equality of polynomial functions.

Fact: Two notions of equality coincide if R an infinite integral domain.

For R finite two notions very different:

$$R = \{ r_1, r_2, \ldots, r_n \},$$
$$Z(x) = (x - r_1)(x - r_2) \cdots (x - r_n).$$

Suppose R is not the zero ring (so $1 \neq 0$). Now

$$Z(x) \neq 0, \quad \text{in } R[x],$$

but

$$\hat{Z}(x) = 0.$$

Polynomials in Several Indeterminates

$R[x]$ a ring.
New indeterminate y.
Get ring $R[x][y]$.
Polynomials in y, coefficients are polynomials in x.
Essentially same ring as $R[y][x]$. (N.B. used $xy = yx$.)
Denote by $R[x, y]$. Elements look like

$$\sum_{i,j=0}^{\infty} r_{ij} x^i y^j,$$

where $r_{ij} \in R$.
Distinguish between *total degree*, $\deg(p)$, *degree in* x, $\deg_x(p)$, and *degree in* y, $\deg_y(p)$.

Can do same for indeterminates $x_1, x_2, \ldots, x_n$.
Power products: expressions

$$x_1^{i_1} \cdots x_n^{i_n}$$

Degree of this is $i_1 + i_2 + \cdots + i_n$.
Notion of degree for polynomials in $R[x_1, x_2, \ldots, x_n]$.
Coefficient of a power product t in a polynomial p: $\mathrm{coeff}(t, p)$.
Convention: if $X = \{ x_1, x_2, \ldots, x_n \}$ write $R[X]$ instead of $R[x_1, x_2, \ldots, x_n]$.

Factorization and Greatest Common Divisors

R a UFD. Then

$$\deg(pq) = \deg(p) + \deg(q), \quad \text{for all } p, q \in R[x].$$

Given non-zero $f \in R[x]$ put $f = ap$ where a is constant (either 1 or is non-invertible) and p has no non-invertible constant factors. Try to express p as:

$$p = hk$$

where $\deg(h) < \deg(p)$, $\deg(k) < \deg(p)$.
Split h, k likewise. Eventually get to

$$p = p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s},$$

where each p_i can't be split up, i.e. it is *irreducible*.

Question: Is this factorization unique?

Answer: Yes if you are careful about what you mean by 'unique'.

Consequence: : R a UFD $\Rightarrow$ $R[x_1, x_2, \ldots, x_n]$ a UFD.

Fact: If R is a UFD then gcd's exist in $R[x]$.

Note: if $p \mid q$ in $R[x]$ then $\deg(p) \leq \deg(q)$.

Fact: Assume $f \neq 0$ or $g \neq 0$. Any gcd h of $f, g \in R[x]$ has maximum possible degree over all common divisors of f, g.

- If p is a common factor, $p \mid h$ so $\deg(p) \leq \deg(h)$.

Question: Given f, g as above with h a gcd. Suppose p is a common divisor of maximum degree how does p relate to h?

Answer: By choice of p we have $\deg(h) \leq \deg(p)$. By above fact $\deg(p) \leq \deg(h)$, i.e., $\deg(p) = \deg(h)$. Since h is a gcd and p a common factor, $p \mid h$. Thus $h = ap$ and so $\deg(a) = 0$, i.e. $a \in R$. Thus p is a gcd except for possibly missing a constant factor.

Fact: Let k be a field and $f, g \in k[x]$. Suppose h is a common factor of highest degree then h is a gcd of f, g. Can make it unique by insisting it is *monic*.

Standard abuse of notation: $\gcd(f, g)$ stands for a gcd of f, g.

Euclid's Algorithm for Univariate Polynomials

Assume coefficients are from a field and $g \neq 0$. Can put

$$f = qg + r, \quad r = 0 \text{ or } \deg(r) < \deg(g).$$

q is *quotient*, r is *remainder*.

Suggested Exercise: Prove that q, r are unique.

Algorithm: Put $r_0 = f$, $r_1 = g$:

$$\begin{aligned} r_0 &= q_1 r_1 + r_2 \\ r_1 &= q_2 r_2 + r_3 \\ r_2 &= q_3 r_3 + r_4 \\ &\vdots \\ r_{s-2} &= q_{s-1} r_{s-1} + r_s \\ r_{s-1} &= q_s r_s + r_{s+1} \end{aligned}$$

where $r_{s+1} = 0$ and $\deg(r_i) < \deg(r_{i-1})$, $1 \leq i \leq s$.

Must eventually have $r_i = 0$ since

$$\deg(r_0) > \deg(r_1) > \ldots > \deg(r_i) > \ldots \geq 0.$$

Rational Coefficients

- Working with fractions $\Rightarrow$ many integer gcd computations.
- Can slow things down.
- Try to use only integer arithmetic.

Fact: If $f, g \in \mathbb{Z}[x]$, $\deg(f) > \deg(g)$ then can find $q, r \in \mathbb{Z}[x]$ s.t.

$$\mathrm{lc}(g)^{\deg(f)-\deg(g)+1} f = qg + r,$$

where $r = 0$ or $\deg(r) < \deg(g)$.

Problem: Coefficients blow up exponentially.

Well Known Example

$$f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5,$$
$$g = 3x^6 + 5x^4 - 4x^2 - 9x + 21.$$

The sequence of remainders obtained by applying the modified algorithm is

$$-15x^4 + 3x^2 - 9,$$

$$15795x^2 + 30375x - 59535,$$

$$1254542875143750x - 1654608338437500,$$

$$12593338795500743100931141992187500.$$

Possible way out: Take out gcd of coefficients at each stage—errm ...
Above board method: Sub-resultant polynomial remainder sequences. OK but compicated.

Extended Euclidean Algorithm for Polynomials

Just like integer case get polys u, v s.t.

$$uf + vg = \gcd(f, g).$$

Moreover can ensure:

$$\begin{array}{lll} u = 0 & \text{or} & \deg(u) < \deg(g) \\ v = 0 & \text{or} & \deg(v) < \deg(f) \end{array}$$

Rational Expressions

- k a field.

$$k(x_1, \ldots, x_n) = \{p/q \mid p, q \in k[x_1, \ldots, x_n] \ \& \ q \neq 0\}.$$

- Equality:

$$p/q = p'/q' \Leftrightarrow pq' - p'q = 0, \quad \text{in } k[x_1, \ldots, x_n].$$

- Define $+$, $*$ by:

$$(p/q) + (p'/q') = (pq' + p'q)/qq',$$
$$(p/q)(p'/q') = pp'/qq'.$$

  Gives us a field.

Caution: Again distinguish between *functions* and elements of $k(x_1, \ldots, x_n)$.

Representation of Polynomials and Rational Expressions

Basic types:

	Dense	Sparse
Recursive		
Distributed		

Recursive Representation

An expression of the isomorphism

$$R[x_1, \ldots, x_n] \cong R[x_1, \ldots, x_{n-1}][x_n].$$

Regard x_n as the main indeterminate.

Example:

$$3xy^2 + 2y^2 - 4x^2y + y - 1$$

represented as

$$(3x + 2)y^2 + (-4x^2 + 1)y + (-1)y^0,$$

y is main indeterminate.

Generally: Use

$$\sum c_i x_n^i$$

each c_i a polynomial represented similarly.

Distributive Representation

Consider power products in given indeterminates e.g.

$$x_1^2 x_3 x_5^7.$$

Pick a total order on power products s.t.

- 1 (i.e. $x_1^0 x_2^0 \cdots x_n^0$) is least,
- each power product has only finitely many others less than it.

Can now write

$$p(x_1, \ldots, x_n) = \sum_{t \leq \bar{t}} c_t t$$

where $c_t \in R$ for each t.

Example suitable ordering:

Total degree then lexicographic.

1. sort according to degree,
2. within each degree use lexicographic ordering: order indeterminates, e.g.

$$x_1 >_L x_2 >_L \cdots >_L x_n$$

then

$$x_1^{i_1} \cdots x_n^{i_n} >_L x_1^{j_1} \cdots x_n^{j_n}$$

if and only if there is a k such that $i_l = j_l$ for $1 \leq l < k$ and $i_k > j_k$.

Dense Representations

Record all coefficients up to highest degree main indeterminate or highest power product.

Example: Recursive representation

$$\sum_{i=0}^{m} c_i x^i \longleftrightarrow (c_0, \ldots, c_m).$$

Example: Distributed representation

$$\sum_{t \leq \bar{t}} c_t t \longleftrightarrow (c_1, c_{t_1}, \ldots, c_{\bar{t}}),$$

where $(\ldots)$ denotes a list or array.

Problem: Can lead to a great deal of wasted space,
Consider $x^{1000} + 1$ or $x^4 y^7 + x + 1$.

Sparse Representations

- Drop all zero coefficients.
- With each non-zero coefficient record corresponding degree or power product.

Example:

$$x^{1000} + 1 \longleftrightarrow ((1, 1000), (1, 0)),$$
$$x^4 y^7 + 2x + 1 \longleftrightarrow ((1, (4, 7)), (2, (1, 0)), (1, (0, 0))).$$

In second example

$$x_1^{e_1} \cdots x_n^{e_n}$$

represented by

$$(e_1, \ldots, e_n).$$

Rational Expressions

- Pair of polynomials $\langle f,g \rangle$
- Numerator in normal form $\Rightarrow \langle f,g \rangle$ in normal form.
- **Dangerous temptation:** Remove $\gcd(f,g)$.
  Consider:
  $$\frac{1-x^n}{1-x} = 1 + x + \cdots + x^{n-1}.$$
  Take e.g. $n = 2^{20}$.
  - L.h.s. needs less than 10 bytes.
  - R.h.s. needs well over a 1,000,000 bytes!
- Nevertheless Axiom does remove $\gcd(f,g)$ automatically, Maple does not.
- Maple uses sum of products representation; very compact but can lead to problems.

Intermediate Expression Swell

Vandermonde determinant

$$V(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}.$$

Basic algebra shows:

$$V(x_1, x_2, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Consider:

$$Z(x_1, x_2, \ldots, x_{n+1}) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_{n+1} \\ x_1^2 & x_2^2 & \ldots & x_{n+1}^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_{n+1}^{n-1} \end{vmatrix}.$$

Obviously:

$$Z(x_1, x_2, \ldots, x_{n+1}) = 0.$$

But expanding along first row:

$$\begin{aligned} Z(x_1, x_2, \ldots, x_{n+1}) &= \sum_{i=1}^{n+1} (-1)^{i+1} V(x_1, \ldots, \hat{x}_i, \ldots, x_{n+1}) \\ &= \sum_{i=1}^{n+1} (-1)^{i+1} \prod_{\substack{1 \le j < k \le n+1 \\ j,k \neq i}} (x_k - x_j), \end{aligned}$$

Perfectly decent sum of products representation.
Expansion leads to $n!$ terms before any cancellation.

Keeping the Data Small: Modular Methods
Gcd of Polynomials in $\mathbb{Z}[x]$

Definition: For $f \in \mathbb{Z}[x]$,

$$f = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$$

define its *content* & *primitive part* by:

$$\operatorname{cont}(f) = \gcd(a_m, a_{m-1}, \ldots, a_0),$$
$$\operatorname{pp}(f) = f/\operatorname{cont}(f).$$

Lemma: (Gauss) For any $f, g \in \mathbb{Z}[x]$ we have $\operatorname{cont}(fg) = \operatorname{cont}(f)\ \operatorname{cont}(g)$ and $\operatorname{pp}(fg) = \operatorname{pp}(f)\ \operatorname{pp}(g)$.

Corollary: : For $f, g \in \mathbb{Z}[x]$

$$\operatorname{cont}(\gcd(f, g)) = \gcd(\operatorname{cont}(f), \operatorname{cont}(g)),$$
$$\operatorname{pp}(\gcd(f, g)) = \gcd(\operatorname{pp}(f), \operatorname{pp}(g)).$$

Conclusion: Can restrict attention to primitive polynomials—gcd also primitive.

Suggested Exercise: Let $f, g \in \mathbb{Z}[x]$ and h be their gcd in $\mathbb{Z}[x]$. Prove that h is also a gcd of f, g in $\mathbb{Q}[x]$.

Useful fact: $\mathrm{lc}(\gcd(f,g)) | \gcd(\mathrm{lc}(f), \mathrm{lc}(g))$.

Equivalantly: If $a \nmid \mathrm{lc}(f)$ or $a \nmid \mathrm{lc}(g)$ then $a \nmid \mathrm{lc}(\gcd(f,g))$.

Definition: Put

$$(f \bmod p) = (a_m \bmod p)x^m + (a_{m-1} \bmod p)x^{m-1} + \cdots + (a_0 \bmod p).$$

Abbreviate $(f \bmod p)$ to f_p. Gives us a function

$$\begin{aligned} \phi : \mathbb{Z}[x] &\to \mathbb{Z}_p[x] \\ f &\mapsto f_p. \end{aligned}$$

$\phi(1) = 1$, $\phi(f+g) = \phi(f) + \phi(g)$, $\phi(fg) = \phi(f)\phi(g)$.

- Example of a *ring homomorphism*.

$$A = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5,$$
$$B = 3x^6 + 5x^4 - 4x^2 - 9x + 21.$$

Put

$$A = PH, \quad B = QH, \quad \text{in } \mathbb{Z}[x],$$

where $H = \gcd(A, B)$. Consider modulo 5;

$$A_5 = P_5 H_5, \quad B_5 = Q_5 H_5, \quad \text{in } \mathbb{Z}_5[x].$$

Direct computation in $\mathbb{Z}_5[x]$ shows:

$$\gcd(A_5, B_5) = 1.$$

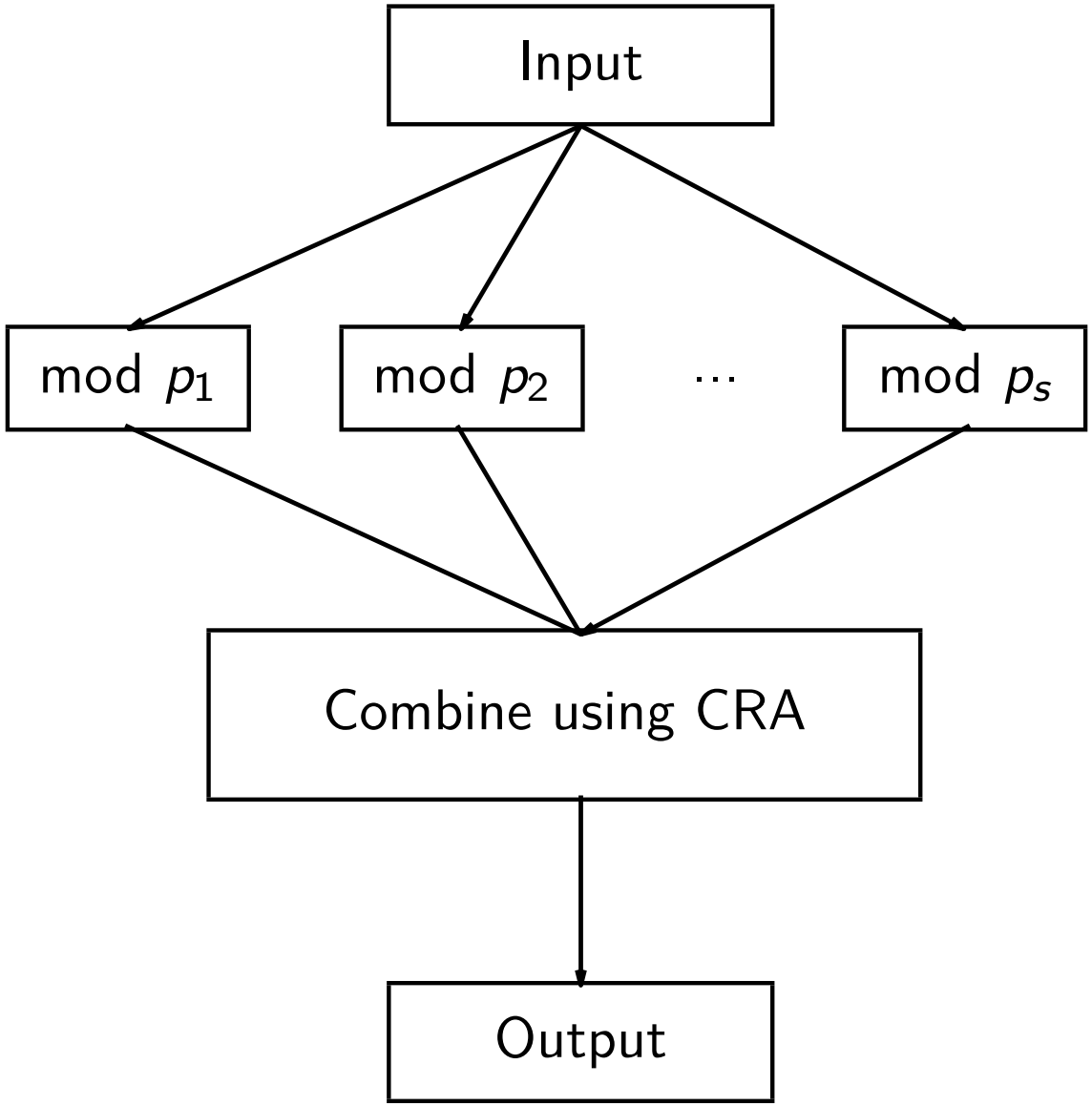
So $H_5 = 1$, more accurately H_5 is a constant. Now

$$\begin{aligned} 5 \not| \operatorname{lc}(A) \ \left[\& \ 5 \not| \operatorname{lc}(B)\right] &\Rightarrow 5 \not| \operatorname{lc}(H) \\ &\Rightarrow \deg(H) = \deg(H_5) \le \deg(\gcd(A_5, B_5)) = 0 \\ &\Rightarrow \deg(H) = 0 \\ &\Rightarrow H \text{ is a constant.} \end{aligned}$$

Thus

$$\gcd(A, B) = 1.$$

General Strategy



Problems to Address

1. How do we combine the various results in the $\mathbb{Z}_{p_i}[x]$ into a single result in $\mathbb{Z}[x]$?
2. Given $A, B \in \mathbb{Z}[x]$ how big can the coefficients of $\gcd(A, B)$ be?
   How do we recover them? (Use symmetric representation of remainders.)
3. Which primes should we choose? Are there any that should be avoided?

Detailed Example

$$A = 3x^4 + 4x^3 - 6x^2 - 3x + 2,$$
$$B = 9x^5 + 21x^4 + 6x^3 + x^2 + x - 2,$$
$$H = \gcd(A, B).$$

Observations:

1. A, B primitive so H primitive.
2. $\deg(H) \leq \min(\deg(A), \deg(B)) = 4$.
3. Easy computation shows $A \not| B$ so $\deg(H) < 4$. Can put

$$H = h_3x^3 + h_2x^2 + h_1x + h_0.$$

Note: Full algorithm does not do this step, only done here to keep number of coefficients down to 4.

Aim: Work modulo p for p a prime (maybe use several p).
Compute

$$F_p = \gcd(A_p, B_p)$$

using Euclid's algorithm in $\mathbb{Z}_p[x]$.

Hope: $F_p = H_p$. *Not* guaranteed.
Sensible to ensure

$$p \not| \operatorname{lc}(A) \text{ or } p \not| \operatorname{lc}(B),$$

so that

$$\deg(F_p) \geq \deg(H_p) = \deg(H).$$

Note: Even if $p \not| \operatorname{lc}(A)$ or $p \not| \operatorname{lc}(B)$ might get

$$\deg(\gcd(A_p, B_p)) > 3$$

which means

$$\gcd(A_p, B_p) \neq H_p.$$

First modulus $p = 2$:

$$A_2 = x^4 + x,$$
$$B_2 = x^5 + x^4 + x^2 + x,$$

Euclid's algorithm in $\mathbb{Z}_2[x]$ gives:

$$\gcd(A_2, B_2) = x^4 + x.$$

Conclusion: Must be something wrong with 2 as a modulus.

Second modulus $p = 3$: No good—divides $\mathrm{lc}(A)$ and $\mathrm{lc}(B)$.

Third modulus $p = 5$:

$$A_5 = 3x^4 + 4x^3 + 4x^2 + 2x + 2,$$
$$B_5 = 4x^5 + x^4 + x^3 + x^2 + x + 3,$$

Get

$$F_5 = \gcd(A_5, B_5) = x^3 + 4x^2 + 2x + 1.$$

No sign of trouble—carry on with hopeful heart.

Test: View F_5 as an element of $\mathbb{Z}[x]$. See if

$$F_5 | A \ \& \ F_5 | B.$$

Test fails: So 5 *might* be a bad choice *or* need more work to recover coefficients of H completely. (At least one of them has been 'collapsed' by taking it modulo 5.)

Fourth modulus $p = 7$:

$$F_7 = \gcd(A_7, B_7) = x^3 + 5x + 4,$$

and $F_7 \not| A$.

Assumption: Both 5 and 7 are good moduli.

Yields: Four pairs of simultaneous congruences:

$$\begin{aligned} h_3 &\equiv 1 \pmod 5, & h_3 &\equiv 1 \pmod 7, \\ h_2 &\equiv 4 \pmod 5, & h_2 &\equiv 0 \pmod 7, \\ h_1 &\equiv 2 \pmod 5, & h_1 &\equiv 5 \pmod 7, \\ h_0 &\equiv 1 \pmod 5, & h_0 &\equiv 4 \pmod 7. \end{aligned}$$

Example: Find all solutions to

$$h_0 \equiv 1 \pmod 5, \qquad h_0 \equiv 4 \pmod 7.$$

First congruence gives:

$$h_0 = 1 + 5q, \quad \text{for } q \in \mathbb{Z}.$$

Substitute into second congruence:

$$5q \equiv 3 \pmod 7.$$

Now

$$3 \cdot 5 - 2 \cdot 7 = 1 \Rightarrow 3 \cdot 5 \equiv 1 \pmod 7$$

So:

$$\begin{aligned} q &\equiv 3 \cdot 3 \pmod 7, \\ &\equiv 2 \pmod 7. \end{aligned}$$

For simultaneous solution take $q = 2 + 7q'$ in $1 + 5q$ to get

$$h_0 = 11 + 35q', \quad \text{for } q' \in \mathbb{Z}$$

i.e.

$$h_0 \equiv 11 \pmod{35}.$$

Solve other pairs of congruences to get:

$$F_{35} = x^3 + 14x^2 + 12x + 11$$

as candidate for H_{35}.

Note: Never did any work modulo 35.

Assumption: Coefficients of H all in range

$$-17 < h \leq 18.$$

Conclusion: Already have H, not just H_{35}.
Simple calculation shows:

$$F_{35} \not| A.$$

Give up?—never!

Crucial observation: When finding gcd's in $\mathbb{Z}_p[x]$ we returned *monic* results.

- In fact any non-zero constant multiple would do just as well but monic is best.
- Assuming p is a good prime, $H_p = lc(H)\gcd(A_p, B_p)$ in $\mathbb{Z}_p[x]$.

Desperate way out: Find $\mathrm{lc}(H)$ and multiply monic gcd's by it.

Much better: Know that

$$\mathrm{lc}(H) \mid c$$

where

$$c = \gcd(\mathrm{lc}(A), \mathrm{lc}(B)) = 3.$$

Take, in $\mathbb{Z}_5[x]$ and $\mathbb{Z}_7[x]$:

$$F_5^* = 3F_5 = 3x^3 + 2x^2 + x + 3,$$
$$F_7^* = 3F_7 = 3x^3 + x + 5.$$

Candidate from F_5^*, F_7^*:

$$F_{35}^* = 3x^3 + 7x^2 + x - 2.$$

Make it primitive—OK already.
Now easy to see

$$F_{35}^* \mid A \ \& \ F_{35}^* \mid B, \quad \text{in } \mathbb{Z}[x],$$

so

$$\gcd(A, B) = F_{35}^*, \quad \text{in } \mathbb{Z}[x].$$

The Chinese Remainder Problem

D a Euclidean domain—i.e. integral domain in which a version of Euclidean Algorithm works.

Given:

1. Remainders $r_1, \ldots, r_n \in D$.
2. Moduli $m_1, \ldots, m_n \in D - \{0\}$ which are pairwise coprime, i.e. $\gcd(m_i, m_j) = 1$ for $i \neq j$.

Problem: Find $r \in D$ such that

$$r \equiv r_i \pmod{m_i}$$

for $1 \leq i \leq n$.

Direct Solution

Let $M_i = m_1 m_2 \cdots m_{i-1} m_{i+1} \cdots m_n$ for $1 \leq i \leq n$.
Find $b_1, b_2, \ldots, b_n$ such that

$$b_i M_i \equiv 1 \pmod{m_i},$$

for $1 \leq i \leq n$ (the b_i exist because $\gcd(M_i, m_i) = 1$).
Then x is a solution to the system

$$\begin{aligned} x &\equiv r_1 \pmod{m_1} \\ x &\equiv r_2 \pmod{m_2} \\ &\vdots \\ x &\equiv r_n \pmod{m_n} \end{aligned}$$

if and only if

$$x \equiv r_1 b_1 M_1 + r_2 b_2 M_2 + \cdots + r_n b_n M_n \pmod{M},$$

where $M = m_1 m_2 \cdots m_n$.

Base Case $n = 2$

$$r \equiv r_1 \pmod{m_1} \tag{1}$$
$$r \equiv r_2 \pmod{m_2} \tag{2}$$

Solutions of (1) have form:

$$r_1 + \sigma m_1.$$

So have to find σ such that:

$$r_1 + \sigma m_1 \equiv r_2 \pmod{m_2}.$$

Use Extended Euclidean Algorithm to find c:

$$cm_1 \equiv 1 \pmod{m_2}.$$

$$\sigma = c(r_2 - r_1) \pmod{m_2}.$$

Thus

$$\begin{aligned} r_1 + \sigma m_1 &\equiv r_1 + c(r_2 - r_1)m_1 \\ &\equiv r_1 + r_2 - r_1 \pmod{m_2}. \end{aligned}$$

Observation: Solution $r = r_1 + \sigma m_1$ is such that the simultaneous congruences

$$x \equiv r_1 \pmod{m_1}$$
$$x \equiv r_2 \pmod{m_2}$$

hold for x if and only if

$$x \equiv r \pmod{m_1 m_2}.$$

General case: Solve first two congruences to obtain r_{12} as answer. General problem now reduces to:

$$x \equiv r_{12} \pmod{m_1 m_2}$$
$$x \equiv r_3 \pmod{m_3}$$
$$\vdots$$

Again have:

$$x \equiv r_i \pmod{m_i}, \quad 1 \leq i \leq n,$$

if and only if

$$x \equiv r \pmod{m_1 m_2 \cdots m_n}.$$

Conclusion

Can work with conveniently sized moduli $m_1, \ldots, m_n$ and then construct result for single large modulus $m_1 m_2 \cdots m_n$.

Theorem: For the case $D = \mathbb{Z}$ the solution r computed by CRA_n or bounded as follows

$$0 \leq r < m_1 m_2 \cdots m_n.$$

Moreover there is exactly one such r.

Theorem: For the case $D = k[x]$ the solution $r(x)$ computed by CRA_n is is either 0 or bounded in degree as follows

$$\deg(r) < \deg(m_1) + \cdots + \deg(m_n).$$

Moreover there is exactly one such $r(x)$.

Suggested Exercise: Prove the claim in the preceding Theorem.

Chinese Remainder Theorem for the Integers

To sum up, stated purely as a theorem we have:

Theorem: Assume $r_1, r_2 \ldots, r_n \in \mathbb{Z}$ and $m_1, m_2, \ldots, m_n \in \mathbb{Z}$ where $m_i > 1$, for $1 \leq i \leq n$, and m_i, m_j are comprime (i.e., $\gcd(m_i, m_j) = 1$) for $1 \leq i < j \leq n$. Then there is an integer x such that

$$\begin{aligned} x &\equiv r_1 \pmod{m_1} \\ x &\equiv r_2 \pmod{m_2} \\ &\vdots \\ x &\equiv r_n \pmod{m_n}. \end{aligned}$$

Moreover setting $M = m_1 m_2 \cdots m_n$ we have that $x + qM$ is also a solution for all $q \in \mathbb{Z}$ and all solutions are of this form.

Integer Case

Choose moduli $m_1, \ldots, m_n$ to be distinct primes:

- Automatically coprime.
- $\mathbb{Z}_p$ a field so in $\mathbb{Z}_p[x]$ gcd's exists and Euclidean Algorithm applies.
  - This is *critical*.
  - p not a prime means $\mathbb{Z}_p$ is not an ID, gcd's need not exist in $\mathbb{Z}_p[x]$.
  - Example: in $\mathbb{Z}_6[x]$ we have

$$3x^d + 1 \mid 2x \quad \& \quad 3x^d + 1 \mid 4x \quad \text{for all } d,$$

since $2x = (3x^d + 1)2x$ and $4x = (3x^d + 1)4x$.

- Use of CRT gives coefficients in range:

$$0 \leq r < M = m_1 m_2 \cdots m_n.$$

But want possibly negative integers.

Shift CRA results to range:

$$-M/2 < r' \leq M/2,$$

where

$$r' = \begin{cases} r, & \text{if } r \leq M/2; \\ r - M, & \text{if } r > M/2. \end{cases}$$

Symmetric representation of remainders.
Can recover R uniquely if $-M/2 < R \leq M/2$.

Conclusion: If trying to recover R with

$$|R| \leq B$$

then choose moduli so that

$$M > 2B.$$

Bound on Coefficients of gcd

Theorem: (Landau-Mignotte Inequality) Let $A = \sum_{i=0}^{m} a_i x^i$ and $B = \sum_{i=0}^{n} b_i x^i$ in $\mathbb{Z}[x]$ and suppose that B is a factor of A. Then

$$\sum_{i=0}^{n} |b_i| \leq 2^n \frac{|b_n|}{|a_m|} \sqrt{\sum_{i=0}^{m} a_i^2}.$$

Corollary: Let $A, B \in \mathbb{Z}[x]$. The absolute value of each coefficient of $\gcd(A, B)$ is bounded by

$$2^{\min(m,n)} \gcd(a_m, b_n) \min\left(\frac{1}{|a_m|} \sqrt{\sum_{i=0}^{m} a_i^2}, \frac{1}{|b_n|} \sqrt{\sum_{i=0}^{n} b_i^2} \right).$$

Conjecture: Coefficients of $\gcd(A, B)$ are no larger in absolute value than the largest absolute value of the coefficients of A or B.

—FALSE—

Bit of a shame really.

Choosing Good Primes

- $A, B \in \mathbb{Z}[x]$, $G = \gcd(A, B)$.
- Choose a prime p s.t. $p \not| \operatorname{lc}(A)$ or $p \not| \operatorname{lc}(B)$ so $p \not| \operatorname{lc}(G)$.

Put

$$A = PG, \quad B = QG,$$

so

$$A_p = P_p G_p, \quad B_p = Q_p G_p.$$

Problem: G_p might not be $\gcd(A_p, B_p)$ in $\mathbb{Z}_p[x]$.

Example: $A = x - 3$, $B = x + 2$, $p = 5$.

$$\begin{array}{rcll} \gcd(A, B) & = & 1, & \text{in } \mathbb{Z}[x], \\ \gcd(A_5, B_5) & = & x + 2, & \text{in } \mathbb{Z}_5[x]. \end{array}$$

Note: We interpret equalities between gcds as being up to an invertible constant multiple.

Lemma: Let $A, B \in \mathbb{Z}[x]$ and p a prime which does not divide both $\mathrm{lc}(A)$, $\mathrm{lc}(B)$. Then

$$\deg(\gcd(A_p, B_p)) \geq \deg(\gcd(A, B)).$$

□

Call a prime p which doesn't work *unlucky*, i.e.

$$\deg(\gcd(A_p, B_p)) > \deg(\gcd(A, B)).$$

Same as

$$\gcd(A_p, B_p) \neq c\gcd(A, B)_p$$

for some constant c.

Note: Could have $\gcd(A_p, B_p) = c\gcd(A, B)_p$ for p dividing both $\mathrm{lc}(A)$, $\mathrm{lc}(B)$. But then we have no reliable way of detecting bad primes.

Question: How many unlucky primes are there?

Very useful tool: introduced by J. Sylvester 19th century.

$$A = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0,$$
$$B = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0,$$

both non-zero. Could have $a_m = 0$ or $b_n = 0$.
The *resultant* of A, B is

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

Have n rows of a-entries, m rows of b-entries, blank spaces 0.

Note: Strictly speaking should wrote $\mathrm{Res}_{m,n}(A,B)$.

Theorem: Suppose that $a_m \neq 0$ or $b_n \neq 0$. Then A and B have a non-constant common factor if and only if $\mathrm{Res}(A, B) = 0$.

Proof: First

Claim: A, B have non-constant common factor iff

$$\psi A = \phi B$$

for some non-zero ϕ and ψ, with

$$\deg(\phi) < m \ \& \ \deg(\psi) < n.$$

Simple proof based on unique factorization.
Now put

$$\phi = \alpha_m x^{m-1} + \cdots + \alpha_1,$$
$$\psi = \beta_n x^{n-1} + \cdots + \beta_1.$$

When can $\psi A = \phi B$?

Equivalent to:

$$\begin{aligned} a_0\beta_1 &= b_0\alpha_1, \\ a_1\beta_1 + a_0\beta_2 &= b_1\alpha_1 + b_0\alpha_2, \\ &\vdots \\ a_m\beta_n &= b_n\alpha_m. \end{aligned}$$

View as set of homogeneous equations in $m + n$ unknowns:

$$\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n.$$

Use determinant condition for existence of non-trivial solution to $MX = \mathbf{0}$. □

Lemma: Let A, B, p, A_p, B_p be as above and put $G = \gcd(A, B)$. Assume that $A_p \neq 0$ and $B_p \neq 0$. If $p \not| \operatorname{Res}(A/G, B/G)$ then

$$\gcd(A_p, B_p) = G_p.$$

Example:

$$\begin{aligned} A &= 3x^4 + 4x^3 - 6x^2 - 3x + 2, \\ B &= 9x^5 + 21x^4 + 6x^3 + x^2 + x - 2, \\ G &= \gcd(A, B) \\ &= 3x^3 + 7x^2 + x - 2. \end{aligned}$$

Thus

$$\begin{aligned} A/G &= x - 1, \\ B/G &= 3x^2 + 1. \end{aligned}$$

So

$$\operatorname{Res}(A/G, B/G) = \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 3 & 0 & 1 \end{vmatrix} = 4$$

$MODGCD(A, B) \mapsto G$

1. $g := \gcd(\mathrm{lc}(A), \mathrm{lc}(B))$;
   $M := 2g\, Landau\_Mignote\_Bound(A, B)$;
2. $p :=$ new prime not dividing g;
3. $C_p := \gcd(A_p, B_p)$ computed in $\mathbb{Z}_p[x]$; (ensure $\mathrm{lc}(C_p) = 1$)
   $G_p := (g \bmod p)C_p$ in $\mathbb{Z}_p[x]$
4. **if** $\deg(G_p) = 0$ **then return** 1 **fi**;
   $P := p$;
   $G := G_p$;
5. **while** $P \leq M$ **do**
      $p :=$ new prime not dividing g;
      $C_p := \gcd(A_p, B_p)$; (ensure $\mathrm{lc}(C_p) = 1$)
      $G_p := (g \bmod p)C_p$;
      if $\deg(G_p) < \deg(G)$ **then goto 4 fi**;
      (all previous primes were unlucky)
      if $\deg(G_p) = \deg(G)$ **then**
         $G := CRA(G, G_p, P, p)$;
         $P := pP$
      fi
   od
6. $H := \mathrm{pp}(G)$;
   if $H \mid A$ **and** $H \mid B$ **then return** H **fi**;
   goto 2 (all the primes were unlucky)

Let

$$\begin{aligned} A &= (x-2)(x+1)(x^3+2x-1) \\ &= x^5 - x^4 - 3x^2 - 3x + 2, \\ B &= (x-2)^2(x+1)^2 \\ &= x^4 - 2x^3 - 3x^2 + 4x + 4. \end{aligned}$$

This yields

$$\begin{aligned} g &= 1, \\ M &= 2 \cdot 1 \cdot 2^4 \cdot 1 \cdot \min(\sqrt{24}, \sqrt{46}) \\ &\leq 160. \end{aligned}$$

Trace of algorithm:

$$
\begin{aligned}
p = 2: \quad & G_2 = x^3 + x, \\
& P = 2, \\
& G = x^3 + x, \\
p = 3: \quad & G_3 = x^2 - x + 1, \text{ so 2 was unlucky;} \\
& P = 3, \\
& G = x^2 - x + 1 \\
p = 5: \quad & G_5 = x^2 - x - 2, \\
& G = x^2 - x - 2, \text{ this is } \gcd(A, B).
\end{aligned}
$$

Note: Algorithm would do 2 more steps to ensure $P > 160$

Polynomial Simplification
Basics of Algebraic Geometry

- k a field,
- $X = \{x_1, \ldots, x_n\}$ indeterminates over k,
- $p_1(x_1, \ldots, x_n), \ldots, p_m(x_1, \ldots, x_n) \in k[X]$.

Definition: The *Variety* corresponding to the polynomials is the set of their common zeros:

$$\mathrm{V}(p_1, \ldots, p_m) = \{ (a_1, \ldots, a_n) \in k^n \mid p_i(a_1, \ldots, a_n) = 0, \text{ for } 1 \leq i \leq n \}.$$

- Subset of k^n (variety depends on k and n).
- Definition makes sense for arbitrary $S \subseteq k[X]$:

$$\mathrm{V}(S) = \{ (a_1, \ldots, a_n) \in k^n \mid p(a_1, \ldots, a_n) = 0, \text{ for all } p \in S \}.$$

Ideals

Take:

$$p_1, \ldots, p_s \in S,$$
$$q_1, \ldots, q_s \in k[X].$$

Put

$$q = q_1 p_1 + \cdots + q_s p_s.$$

Obviously

$$q(a_1, \ldots, a_n) = 0, \quad \text{for all } (a_1, \ldots, a_n) \in \mathrm{V}(S).$$

Thus

$$\mathrm{V}(S \cup \{q\}) = \mathrm{V}(S).$$

Can add any set of polynomials like q to S without changing the variety.

Definition: The *ideal* of $k[X]$ generated by S, denoted by (S), is:

$$(S) = \{\, q_1 p_1 + \cdots + q_s p_s \mid s \geq 1, q_i \in k[X], p_i \in S, \text{ for } 1 \leq i \leq s \,\}.$$

Have

$$\mathrm{V}(S) = \mathrm{V}((S)).$$

Say that S is a *basis* of ideal I if $I = (S)$. (I is *generated* by S.)

Note: Bases *not* unique.

Note: Exactly the same definition of ideal applies to arbitrary commutative rings.

Abstract definition: I is an ideal if and only if

1. $I \neq \emptyset$,
2. $p_1, p_2 \in I \Rightarrow p_1 q, p_1 - p_2 \in I$ for all $q \in k[X]$.

Fact: If $S_1 \subseteq S_2$ then $(S_1) \subseteq (S_2)$.

Fact: If I is an ideal and $p_1, \ldots, p_s \in I$, $q_1, \ldots, q_s \in k[X]$ then $q_1 p_1 + \cdots + q_s p_s \in I$.

Fact: If I is an ideal and $S \subseteq I$ then $(S) \subseteq I$.

$S \subseteq k[x, y]$ with elements

$$p_1 = x^2y + x - 1,$$
$$p_2 = xy^2 + y - 1.$$

Then (S) contains

$$(2x + 3y^2)p_1 = 3x^2y^3 + 2x^3y + 3xy^2 + 2x^2 - 3y^2 - 2,$$
$$yp_1 - xp_2 = x - y,$$

and infinitely more.

Consider

$$p_1 = x + y - 2z - 1,$$
$$p_2 = 2x - 3y - z + 2,$$
$$p_3 = x - y + z,$$

from $\mathbb{Q}[x, y, z]$ and let $I = (p_1, p_2, p_3)$. Now

$$p_4 = p_2 - 2p_1 = -5y + 3z + 4 \in I$$

Therefore

$$p_5 = p_3 - p_1 - 2/5p_4 = 9/5z - 3/5 \in I$$

Thus

$$(p_1, p_4, p_5) \subseteq I.$$

Easily $p_2, p_3 \in (p_1, p_4, p_5)$ so

$$I = (p_1, p_4, p_5).$$

Thus

$$\begin{aligned} V(I) &= V(p_1, p_4, p_5) \\ &= V(x + y - 2z - 1, -5y + 3z + 4, 9/5z - 3/5). \end{aligned}$$

Final set of equations is in triangular form so very easy to solve.

Major Problem

Question: Does every ideal have a *finite* basis?.

Geometric significance: Given figures in n dimensional space defined by *infinitely* many polynomial equations. Are there *finitely* many equations that define precisely the same figures?

$|X| = 1$: Yes—easy (follows from Euclidean Algorithm).

$|X| = 2$: Yes—long & complicated proof by Gordan (the 'King of the invariants').

$|X|$ arbitrary: Yes—Hilbert's Basis Theorem (1888) very short proof!

Theorem: [Hilbert's Basis Theorem, (1888)] Every ideal of $k[X]$ has a finite basis.

Method of proof: non-constructive.

Gordan's reaction: 'Das ist nicht Mathematik. Das ist Theologie'. Not just sour grapes—fairly typical at the time.

Later on: Hilbert produced constructive proof based on earlier non-constructive one.

Can view V as a function

$$\text{Ideals} \to \text{Varieties}.$$

Have obvious function I in opposite direction:

$$\text{Varieties} \to \text{Ideals}$$

assigns to variety V the ideal

$$\mathrm{I}(V) = \{\, p \mid p \in k[X] \;\&\; p(a_1, \ldots, a_n) = 0, \text{ for all } (a_1, \ldots, a_n) \in V \,\}.$$

Questions:

1. is $I = \mathrm{I}\,\mathrm{V}(I)$ for an arbitrary ideal I of $k[X]$?
2. is $V = \mathrm{V}\,\mathrm{I}(V)$ for an arbitrary variety V of k^n?

Easily:

1. $I \subseteq \mathrm{I}\,\mathrm{V}(I)$ for all ideals I of $k[X]$,
2. $V \subseteq \mathrm{V}\,\mathrm{I}(V)$ for all varieties V of k^n,

In fact always have

$$V = \mathrm{V}\,\mathrm{I}(V)$$

But can have

$$I \neq \mathrm{I}\,\mathrm{V}(I),$$

e.g. take $V = \mathrm{V}(p(x)^2)$, $p(x)$ non-constant.

Definition: k is *algebraically closed* if every non-constant $p \in k[x]$ has a root in k

Example: $\mathbb{C}$, field of complex numbers.

Assumption: from now on k is algebraically closed.

Theorem: [Hilbert's Nullstellensatz, (1893)] Let I be an ideal of $k[X]$ and q a polynomial of $k[X]$ which is zero at all points of $\mathrm{V}(I)$, i.e. $q \in \mathrm{IV}(I)$. Then $q^s \in I$ for some integer $s > 0$.

Concrete form: If $q, p_1, \ldots, p_m \in k[X]$ and q vanishes whenever $p_1, \ldots, p_m$ do then there exist $s > 0$ and $q_1, \ldots, q_m \in k[X]$ such that

$$q^s = q_1 p_1 + \cdots + q_m p_m.$$

Equivalent form: $\mathrm{V}(I) = \emptyset$ if and only if $1 \in I$ (i.e. $I = k[X]$).

Concrete form: A simultaneous system of polynomial equations:

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0 \\ p_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ p_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

does *not* have a simultaneous solution if and only if

$$1 = q_1 p_1 + \cdots + q_m p_m$$

for some $q_1, \ldots, q_m \in k[X]$.

Note: Nullstellensatz definitely *false* if k not algebraically closed: consider $p = x^2 + 1 \in \mathbb{R}[x]$.

Gröbner Bases

Polynomial Ideal Membership: Given polynomials $q, p_1, \ldots, p_m \in k[X]$ is

$$q \in (p_1, \ldots, p_m)?$$

Answer: Compute *Gröbner basis* G of $(p_1, \ldots, p_m)$. Return *yes* if q reduces to 0 w.r.t. G else *no*.

Observation: $p_1, \ldots, p_m$ have simultaneous solution if and only if G does not contain a non-zero constant.

Fact: Gröbner basis of an ideal is a canonical form for it provided basis is *normed* and *reduced*.

$[X]$ denotes set of all power products in indeterminates of X:

$$x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n}.$$

Example of a *monoid*.
Take

$$F = \{ p_1, \ldots, p_m \}.$$

By definition

$$q \in (F)$$

if and only if

$$q = p_1 q_1 + \cdots + p_m q_m,$$

for some $q_1, \ldots, q_m \in k[X]$.
Equivalently

$$q = c_1 f_1 s_1 + \cdots + c_r f_r s_r,$$

where $c_i \in k$, $f_i \in F$, $s_i \in [X]$ for $1 \leq i \leq r$.

Definition: Write

$$g \to_F h$$

to mean

$$h = g - cfs$$

for some $c \in k$, $f \in F$ and $s \in [X]$.

- Like a division step.

Say that g *reduces* to h w.r.t. F.

Rephrasing: $q \in (F)$ if and only if there is sequence

$$q = q_1 \to_F q_2 \to_F \cdots \to_F q_r = 0.$$

Problem: Infinitely many choices at each reduction step.

Possible solution: Introduce suitable order on power products & avoid reductions that introduce bigger power products than previously seen.
Try to 'squeeze' q down to 0.

In trying to find a reduction

$$q_i \to_F q_{i+1}$$

aim to kill at least one power product of q_i (might introduce some new smaller ones into q_{i+1}).

1. Pick victim power product v from q_i (could always take largest).
2. Find some $f \in F$ whose leading power product u divides v, i.e.

$$v = ut, \quad \text{for some } t \in [X].$$

3. Now

$$q_i \to_F q_i - \frac{\operatorname{coeff}(v, q_i)}{\operatorname{lc}(f)} ft$$

Assumption: Multiplication by power products respects order ($u > v \Rightarrow ut > vt$).
Call a sequence of such reductions *restricted* (previous type called *unrestricted*).

Note: At each step have only finitely many possible reductions (assuming ordering is reasonable).
Moreover there are no infinite chains of reductions (*not* obvious).

Algorithm: Given q, F construct the finite tree. If any leaf holds 0 then $q \in (F)$ else $q \notin (F)$.

Take

$$\begin{aligned} p_1 &= y - 1, \\ p_2 &= x, \\ q &= xy + x. \end{aligned}$$

Then, for any sensible order:

$$\begin{aligned} xy + x &\to_{p_1} (xy + x) - x(y - 1) \\ &= 2x \\ &\to_{p_2} 2x - 2x \\ &= 0 \end{aligned}$$

Conclusion: $xy + x \in (y - 1, x)$.

Take

$$p_1 = y + 1,$$
$$p_2 = xy,$$
$$q = x.$$

Order: any admissible (e.g., lexicographic with $x <_L y$).

Clearly: no reductions apply to q.

Conclusion: $q \notin (p_1, p_2)$.
Alas

$$q = xp_1 - p_2$$

so

$$q \in (p_1, p_2).$$

Source of error: We really *do* have that if q reduces to 0 (by restricted sequence) then $q \in (F)$. *But* we did not prove the converse. (Of course OK if unrestricted reductions allowed.)

Basic problem: Unrestricted reduction sequence can introduce power products of any order which are cancelled later on.

Idea: Change basis for ideal to avoid cancellation difficulty.

Want: New finite basis G of (F) such that if q reduces to 0 by unrestricted reduction sequence (w.r.t. F or G) then same happens via restricted reduction sequence (w.r.t. G).

Lemma: Suppose $f \to_F g$ then $f \in (F)$ if and only if $g \in (F)$.

Corollary: Suppose $f_1 \to_F f_2 \to_F \cdots \to_F f_n$ then $f_1 \in (F)$ if and only if $f_n \in (F)$.

Observation: $0 \in (F)$ so if we get to 0 this is a proof that the starting polynomial is in (F).

Question: When can the restricted reductions idea go wrong?

1. $f_1 \notin (F)$: all reduction sequences will stop at a non-zero polynomial. This is correct.
2. $f_1 \in (F)$:
   i. Reduce to 0, this is fine it proves that $f_1 \in (F)$.
   ii. Reductions stop with g and $g \neq 0$; this is wrong, it implies that $f_1 \notin (F)$ in our algorithm.

Bright idea: Suppose basis G of (F) has the property that

$$f \in (F) \ \& \ f \neq 0 \Rightarrow \mathrm{lpp}(p) \mid \mathrm{lpp}(f), \quad \text{for some } p \in G.$$

Then 2.ii cannot happen.

Question: Does such a G exist?

Answer: Yes.

Sketch proof:

- Consider ideal L generated by all leading power products of all elements of (F):

$$L = \big(\{\mathrm{lpp}(f) \mid f \in (F), f \neq 0\}\big).$$

- Take finite basis M of L—existence guaranteed by Hilbert's Basis Theorem (can assume M consists of power products).
- Now take any finite subset G of (F) whose leading power products include all those of M.
- G is desired basis.

Given that G exists how can we compute it?

Pseudo-Algorithm: $Gr(F) \mapsto G$

1. $G := F$;
2. Try to create new element $p \in (F)$ such that
$$\mathrm{lpp}(g) \not| \mathrm{lpp}(p)$$
for all $g \in G$.
3. **if** (no such p can be created) **then return** G
else $G := G \cup \{ p \}$; **goto** 2
fi

All elements of (G) have form:

$$c_1 f_1 s_1 + \cdots + c_r f_r s_r,$$

where $c_i \in k$, $f_i \in G$, $s_i \in [X]$.
If all $f_i = f$, say then nothing new. Must use at least take two different f_i:

$$c_1 f_1 s_1 + c_2 f_2 s_2.$$

Bit of thought leads to:

$$\begin{aligned}\operatorname{spol}(f,g) = &\frac{1}{\operatorname{lc}(f_1)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f_1), \operatorname{lpp}(f_2))}{\operatorname{lpp}(f_1)} \cdot f_1 \\ &- \frac{1}{\operatorname{lc}(f_2)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f_1), \operatorname{lpp}(f_2))}{\operatorname{lpp}(f_2)} \cdot f_2.\end{aligned}$$

Reduce w.r.t. G as far as possible to get h. If $h = 0$ then nothing new. Else put h into basis.

Fact: Eventually all such polynomials reduce to 0, i.e. process stops.

Definition and Characterization of Gröbner Bases

Admissible ordering on $[X]$: must satisfy

1. $1 < t$, for all $t \in [X] - \{1\}$.
2. $s < t \Rightarrow su < tu$, for all $s, t, u \in [X]$.

Lemma: Let $\leq$ be any admissible ordering. Then

1. $s \mid t \Rightarrow s \leq t$ for all $s, t \in [X]$.
2. There are no infinite decreasing sequences (Noetherianity). □

Lexicographic Ordering: order indeterminates, e.g.

$$x_1 >_L x_2 >_L \cdots >_L x_n.$$

Then

$$x_1^{i_1} \cdots x_n^{i_n} >_L x_1^{j_1} \cdots x_n^{j_n}$$

iff there is an r s.t. $i_l = j_l$ for $1 \leq l < r$ and $i_r > j_r$.

Graduated Lexicographic Ordering: put

$$s <_G t \Longleftrightarrow \deg(s) < \deg(t) \text{ or}$$
$$\deg(s) = \deg(t) \text{ and } s <_L t.$$

Second ordering also called *total degree then lexicographic*.

Notation

Given: $f \in k[X] - \{0\}$.

Leading power product:

$$\mathrm{lpp}(f) = \max_{<}\{\, t \in [X] \mid \mathrm{coeff}(t, f) \neq 0 \,\}.$$

Leading coefficient:

$$\mathrm{lc}(f) = \mathrm{coeff}(\mathrm{lpp}(f), f).$$

Initial term: (or initial monomial)

$$\mathrm{in}(f) = \mathrm{lc}(f)\,\mathrm{lpp}(f).$$

Extend to sets: For $F \subseteq k[X]$ put

$$\mathrm{lpp}(F) = \{\, \mathrm{lpp}(f) \mid f \in F,\ f \neq 0 \,\},$$
$$\mathrm{in}(F) = \{\, \mathrm{in}(f) \mid f \in F,\ f \neq 0 \,\}.$$

Definition: J a non-zero ideal of $k[X]$. Finite subset G of $J - \{0\}$ is a *Gröbner basis* for J if

$$\big(\mathrm{in}(G)\big) = \big(\mathrm{in}(J)\big).$$

Just a compressed way of saying:

> *For all $f \in J$ with $f \neq 0$ there is a $g \in G$ such that $\mathrm{lpp}(g) \mid \mathrm{lpp}(f)$.*

Observations:

1. Finite subset $G \subseteq J - \{0\}$ is a Gröbner basis for J iff $\mathrm{lpp}(J) = \mathrm{lpp}(G) \cdot [X]$.
2. Every ideal J of $k[X]$ has a Gröbner basis.
3. If G is a Gröbner basis for J and $f \in J$ then $G \cup \{f\}$ is a Gröbner basis for J.
4. G is a Gröbner basis for J iff $(\mathrm{lpp}(G)) = (\mathrm{lpp}(J))$.
5. Not every basis for an ideal is a Gröbner basis.
6. Any set of monomials $\{ c_1 t_1, \ldots, c_m t_m \}$ is a Gröbner basis for the ideal it generates.

Definition: Reduction relation $\to_F$ as before.
$\to_F^*$ means apply $\to_F$ zero or more times.

Theorem: $\to_F^*$ always terminates. □

Example:

$$f_1 = x^2y^2 + y - 1,$$

$$f_2 = x^2y + x,$$

$$F = \{ f_1, f_2 \} \subseteq \mathbb{Q}[x, y].$$

Order is lexicographic with $x <_L y$:

$$2\underbrace{x^2y^3}_{s} + x^2y + 1 \to_{f_1} (2x^2y^3 + x^2y + 1) - 2yf_1$$

$$= -2y^2 + \underbrace{x^2y}_{s} + 2y + 1$$

$$\to_{f_2} (-2y^2 + x^2y + 2y + 1) - 1 \cdot 1 \cdot f_2$$

$$= -2y^2 + 2y - x + 1.$$

Definition: Given $f, g \in k[X]$.

$$\operatorname{spol}(f,g) = \frac{1}{\operatorname{lc}(f)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f), \operatorname{lpp}(g))}{\operatorname{lpp}(f)} \cdot f$$
$$- \frac{1}{\operatorname{lc}(g)} \cdot \frac{\operatorname{lcm}(\operatorname{lpp}(f), \operatorname{lpp}(g))}{\operatorname{lpp}(g)} \cdot g.$$

Example:

$$f = 2x^2y + 3x^2 + 1,$$
$$g = 3xy^2 - 2x.$$

Then

$$\begin{array}{ccc} & x^2y^2 & \\ \swarrow f & & \searrow g \\ (1/2)y(3x^2+1) & & (1/3)x(-2x) \end{array}$$

Difference of new polynomials is $\operatorname{spol}(f,g)$.

Computation of Gröbner Bases

GRÖBNER_BASIS$(F) \mapsto G$
(F and G are finite sets of polynomials, $(F) = (G)$ and G is a Gröbner basis for (F).)
$G := F$;
while not all S-polys of G have been considered **do**
  choose a new $\mathrm{spol}(f, g)$;
  compute a normal form h of it w.r.t. G;
  if $h \neq 0$ **then** $G := G \cup \{h\}$ **fi**
od

GRÖBNER_BASIS$(F) \mapsto G$
(F and G are finite sets of polynomials, $(F) = (G)$ and G is a Gröbner basis for (F).)
let $F = \{ f_1, f_2, \ldots, f_m \}$;
$P := \{(f_i, f_j) \mid 1 \leq i < j \leq m\}$;
$G := F$;
while $P \neq \emptyset$ **do**
  remove a pair (f, g) from P;
  compute a normal form h of $\mathrm{spol}(f, g)$ w.r.t. G;
  if $h \neq 0$ **then**
    $P := P \cup \{(h, p) \mid p \in G\}$;
    $G := G \cup \{h\}$
  fi
od

Example: $f_1 = x^2y^2 + y - 1$, $f_2 = x^2y + x$.
Order: Lexicographic with $x <_L y$.

1. $\mathrm{spol}(f_1, f_2) = f_1 - yf_2 = -xy + y - 1$. Irreducible.

$$f_3 = -xy + y - 1, \quad G = \{f_1, f_2, f_3\}.$$

2. $\mathrm{spol}(f_2, f_3) = f_2 + xf_3 = xy \to_{f_3} y - 1$. Put

$$f_4 = y - 1, \quad G = \{f_1, f_2, f_3, f_4\}.$$

3. $\mathrm{spol}(f_3, f_4) = f_3 + xf_4 = -x + y - 1 \to_{f_4} -x$. Put

$$f_5 = -x, \quad G = \{f_1, f_2, f_3, f_4, f_5\}.$$

4. $\mathrm{spol}(f_1, f_3) = f_1 + xyf_3 = xy^2 - xy + y - 1 \to_G 0$.
5. $\mathrm{spol}(f_2, f_4) = f_2 - x^2f_4 = x^2 + x \to_G 0$.

Output:

$$G = \{x^2y^2 + y - 1,\ x^2y + x,\ -xy + y - 1,\ y - 1,\ -x\}.$$

In fact $\{x,\ y - 1\}$ is a Gröbner basis for $\{f_1, f_2\}$.

Theorem: Let G be a Gröbner basis for an ideal I of $K[X]$. Let $g, h \in G$ with $g \neq h$. Then

1. If $\mathrm{lpp}(g) \mid \mathrm{lpp}(h)$ then $G' = G - \{h\}$ is also a Gröbner basis for I.
2. If $h \rightarrow_{G-\{h\}} h'$ then $G' = (G - \{h\}) \cup \{h'\}$ is also a Gröbner basis for I. □

Remarks:

- G is *minimal* iff $\mathrm{lpp}(g) \nmid \mathrm{lpp}(h)$ for all $g \neq h \in G$.
- G is *reduced* iff for all $h \neq g \in G$, h cannot be reduced by g.
- G is a *normed* basis if $\mathrm{lc}(f) = 1$ for all $f \in G$.

Theorem: A normed reduced basis for an ideal I is unique. □

Applications of Gröbner Bases

Very many applications exist.

Ideal Membership: Given an ideal $I = (f_1, \ldots, f_s)$ and a polynomial f is $f \in I$?

Solution: Compute a Gröbner basis G for I. Then

$$f \in I \iff f \to_G^* 0.$$

Solution of Equations: Hilbert's Nullstellensatz says:

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0 \\ p_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ p_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

does *not* have a simultaneous solution (over $\mathbb{C}$) iff

$$1 \in (p_1, p_2, \ldots, p_m).$$

For any ideal I

$$1 \in I \Leftrightarrow a \in G, \quad a \in \mathbb{C} - \{0\},\ G \text{ any Gröbner basis of } I,$$
$$\Leftrightarrow N = \{1\}, \quad N \text{ the } \textit{normed} \text{ Gröbner basis of } I.$$

Less obvious:

Theorem: A system of polynomial equations has finitely many solutions over $\mathbb{C}$ if and only if each indeterminate appears in the form cx^d, where c is a constant, as the initial term of one of the members of the reduced Gröbner basis of polynomials where the basis is computed w.r.t. a lexicographic ordering.

Diagonalization of system—cf. Gaussian elimination for linear equations.

Suppose $x_1 >_L x_2 >_L \cdots >_L x_n$. Theorem says there are finitely many solutions if and only if lexicographic Gröbner basis has a subset that looks like:

$$\begin{aligned} c_1 x_1^{d_1} + p_1(x_1, x_2, \ldots, x_n) \\ c_2 x_2^{d_2} + p_2(x_2, \ldots, x_n) \\ \ddots \\ c_n x_n^{d_n} + p_n(x_n) \end{aligned}$$

where $c_i \neq 0$, for $1 \leq i \leq n$.

Note: This includes possibility that $d_i = 0$. Then ith polynomial is just $c_i \neq 0$ so system has 0 solutions.

Example:

$$f = 3x^2y + 2xy + y + 9x^2 + 5x - 3,$$
$$g = 2x^3y - xy - y + 6x^3 - 2x^2 - 3x + 3,$$
$$h = x^3y + x^2y + 3x^3 + 2x^2.$$

Use lexicographic order with $x >_L y$, basis:

$$\{\, 21 - 16y - 3y^2 + 2y^3,\ 8x - 2y^2 + 5y + 3 \,\}.$$

Solutions:

$$(-3/8 - 5/8a + 1/4a^2, a)$$

with a a root of $21 - 16y - 3y^2 + 2y^3$.

Example:

$$f = (y - 1)x + y - 1,$$
$$g = y^2 - 1.$$

Use lexicographic order with $x >_L y$, basis:

$$\{\, y^2 - 1,\ xy - x + y - 1 \,\}.$$

Infinitely many solutions.

Cost of Method

Question: If $q \in (p_1, \ldots, p_m)$ what degrees can occur for $q_1, \ldots, q_m$ of minimal degree so that $q = q_1 p_1 + \cdots + q_m p_m$?

Fact: For infinite fields, degree is at most double exponential in number n of indeterminates. Moreover this is necessary for certain examples.
Given Gröbner basis for $(p_1, \ldots, p_m)$ then, assuming $q \in (p_1, \ldots, p_m)$, can find $q_1, \ldots, q_m$ for little extra cost.

Conclusion: Double exponential space lower bound applies to construction of Gröbner bases.

But: Algorithm runs fine with very many examples of interest.

Moral: Only use it if you really need to—can't use it as a matter of course.

Suggested Exercise: 6.5.

Real Roots of Polynomials

Given: $p(x) \in \mathbb{Q}[x]$, $p(x) \neq 0$.

Find: All real roots of $p(x)$.

Want: Absolute reliability—rules out Newton-Raphson etc.

Fourier's approach:

Isolation: find open intervals s.t. each one contains exactly one real root of p and each real root of p is contained in an interval.

Approximation: shrink each interval to approximate root it contains to desired degree of accuracy.

Real Root Approximation

Note:

- p has no root in $(a, b) \Rightarrow p$ doesn't change sign in (a, b).
- Converse false: e.g., $p = x^2$, $(a, b) = (-1, 1)$.

Theorem: If p is square free then it has a root in (a, b) iff it changes sign.

Computing square-free part:

$$p(x) = \prod_{i=1}^{r} p_i(x)^i, \quad \text{each } p_i \text{ square free, possibly some } p_i = 1.$$

$$p'(x) = \sum_{i=1}^{r} i p_i(x)^{i-1} p_i'(x) \prod_{j \neq i} p_j(x)^j.$$

$$r(x) = \gcd(p(x), p'(x)) = \prod_{i=2}^{r} p_i(x)^{i-1}.$$

$$p(x)/r(x) = \prod_{i=1}^{r} p_i(x) = p/\gcd(p, p').$$

Subtlety: Method used is over $\mathbb{Q}[x]$. *But* want square-free part of p as element of $\mathbb{R}[x]$.

Theorem: Let D, D' be integral domains with D a subdomain of D'. Suppose $f, g \in D[x]$ have a non-constant common factor in $D'[x]$, then they have a non-constant common factor in $D[x]$.

More straightforwardly: the Euclidean Algorithm shows that the result of computing $r(x)$ is unchanged if we work over $\mathbb{Q}$ or $\mathbb{R}$.

Reason:

- The algorithm uses the coefficients of the input polynomials and just adds, subtracts, multiplies (or divides) by them.
- No other field elements are used.
- So if the coefficients come from a field k the whole computation stays in k.

Conclusion: If a property can be defined in terms of gcd and standard ring operations it is very *robust*, i.e., if k is a subfield of k' and $f \in k[x]$ then f has the property in $k'[x]$ if and only if it has it in $k[x]$.

$APPROX(p(x), a, b, \epsilon) \mapsto A$
(A is either the exact root of p contained in (a, b) or an interval (c, d) which isolates the same root and satisfies $d - c < \epsilon$.)

$q(x) := SQFPART(p)$;
if $q(a) = 0$ **then** $q(x) := q(x)/(x - a)$ **fi**;
if $q(b) = 0$ **then** $q(x) := q(x)/(x - b)$ **fi**;
$c := a$;
$d := b$;
$m := (c + d)/2$;
while $d - c \geq \epsilon$ **do**
  if $q(m) = 0$ **then return** m **fi**;
  if $sign(q(c)) \neq sign(q(m))$ **then**
    $d := m$; $m := (c + d)/2$
  else $c := m$; $m := (c + d)/2$
  fi
od;
$A := (c, d)$

Real Root Isolation

$ISOL(p(x), a, b) \mapsto [E, A]$

($p(x)$ is square free. E is a list of (some of the) exact roots of $p(x)$ which lie in (a, b) and A is a list of isolating intervals for the rest of the roots of $p(x)$ in (a, b).)

1. $E := [\,]$; $A := [\,]$;
$r := RCOUNT(p(x), a, b)$;
2. **if** $r = 0$ **then return** $[[\,], [\,]]$
elif $r = 1$ **then return** $[[\,], [(a, b)]]$
fi;
3. $W := [[a, b, r]]$; (to be explored further)
4. **while** $W \neq [\,]$ **do**
remove the first element $[c, d, r]$ from W;
$m := (c + d)/2$;
if $p(m) = 0$ **then**
$E := [op(E), m]$;
$p(x) := p(x)/(x - m)$
fi;
$r := RCOUNT(p(x), a, m)$;
if $r = 1$ **then** $A := [op(A), (a, m)]$
elif $r > 1$ **then** $W := [op(W), [a, m, r]]$
fi;
$r := RCOUNT(p, m, b)$;
if $r = 1$ **then** $A := [op(A), (m, b)]$
elif $r > 1$ **then** $W := [op(W), [m, b, r]]$
fi;
od

Bounding the Real Roots

Two theorems by Cauchy.

Theorem: Let

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$$

be a polynomial with complex coefficients where $m \geq 1$ and $a_m \neq 0$. Then any root α of p satisfies

$$|\alpha| < 1 + \frac{\max\{\,|a_0|, \ldots, |a_{m-1}|\,\}}{|a_m|}.$$

Theorem: Let

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$$

be a polynomial with real coefficients where $m \geq 1$, $a_m > 0$ and which has $\lambda > 0$ strictly negative coefficients. Put

$$B = \max\left\{ \left| \lambda \frac{a_{m-i}}{a_m} \right|^{1/i} \;\middle|\; 1 \leq i \leq m \;\&\; a_{m-i} < 0 \right\}.$$

Then every positive real root of p is no larger than B.

Counting Real Roots of Univariate Polynomials

Sequence	Variation
$2, \quad -1, \quad -2, \quad 3, \quad 3$	2
$2, \quad 0, \quad -1, \quad 0, \quad 0, \quad -2, \quad 3, \quad 3, \quad$	2
$-1, \quad -10, \quad 0, \quad -3, \quad 0, \quad 1, \quad 2, \quad 0, \quad 3$	1
$0, \quad 3, \quad -1, \quad 2, \quad 0, \quad -4, \quad 5, \quad 0, \quad -6$	5

$$S = x^3 - 7x + 7,\ 3x^2 - 7,\ 2x - 3,\ 1.$$

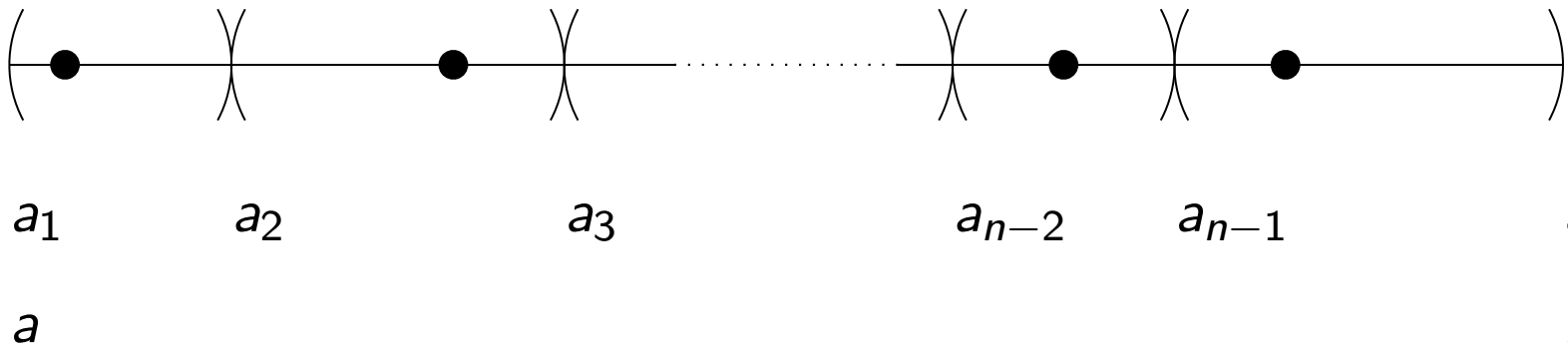
x	Sequence	$V_S(x)$
-1	$13, \quad -4, \quad -5, \quad 1$	2
0	$7, \quad -7, \quad -3, \quad 1$	2
$1/2$	$29/8, \quad -25/4, \quad -2, \quad 1$	2
1	$1, \quad -4, \quad -1, \quad 1$	2
$3/2$	$-1/8, \quad -1/4, \quad 0, \quad 1$	1
2	$1, \quad 5, \quad 1, \quad 1$	0

Note: $V_S(x)$ cannot change as x varies unless x passes through a root of some polynomial of S.

Suppose

$$S = p_0(x), p_1(x), \ldots, p_m(x)$$

all polynomials non-zero. Then (a, b) has only finitely many numbers that are roots of some $p_i(x)$. So can cut up as:



$$V_S(a) - V_S(b) = \sum_{i=1}^{n-1} V_S(a_i) - V_S(a_{i+1}).$$

Suppose

$$V_S(a_i) - V_S(a_{i+1}) \geq 1$$

when (a_i, a_{i+1}) contains a root of some $p_i(x)$. Then

$$\text{number of such roots in } (a, b) \leq V_S(a) - V_S(b).$$

Given $p(x)$ suppose can find sequence S s.t.

$$V_S(a_i) - V_S(a_{i+1}) = \begin{cases} 1, & \text{if } p(x) \text{ has a root in } (a_i, a_{i+1}); \\ 0, & \text{otherwise.} \end{cases}$$

Conclusion:

$$\text{number of roots of } p(x) \text{ in } (a, b) = V_S(a) - V_S(b).$$

Sturm Sequences

Assume $p(x)$ square free.

$$\mathrm{Sturm}(p) = p_0(x), p_1(x), \ldots, p_n(x)$$

defined by

$$\begin{aligned} p_0(x) &= p(x), \\ p_1(x) &= p'(x), \\ p_i(x) &= -\text{remainder}(p_{i-2}(x), p_{i-1}(x)), \quad \text{for } i \geq 2. \end{aligned}$$

Euclidean algorithm with negative remainders.
Stop when non-zero constant reached

$$\begin{aligned} p_0(x) &= p_1(x)q_1(x) - p_2(x) \\ p_1(x) &= p_2(x)q_2(x) - p_3(x) \\ &\vdots \\ p_{n-2}(x) &= p_{n-1}(x)q_{n-1} - p_n(x) \end{aligned}$$

So

$$p_n(x) = \gcd(p_0(x), p_1(x) = \text{non-zero constant}.$$

Properties of Sturm Sequences

1. If α a real root of $p(x)$ then for a sufficiently small ϵ, $p(x)$ and $p'(x)$ have opposite sign for all $x \in (\alpha - \epsilon, \alpha)$ and the same sign in $(\alpha, \alpha + \epsilon)$.
   (True even if $p(x)$ not square free.)
2. Two consecutive elements of the sequence cannot vanish at the same point.
3. If $p_i(\alpha) = 0$ for some i with $0 < i < n$ then $p_{i-1}(\alpha)$ and $p_{i+1}(\alpha)$ have opposite signs.

Sturm's Theorem (1835): Let $p(x)$ be square free and let a, b be two real numbers. Then the number of real roots of $p(x)$ in the interval $(a, b]$ is $V_S(a) - V_S(b)$ where $S = \mathrm{Sturm}(p)$.

Counting all Real Roots

Given $p(x) \neq 0$.
$(-a, a)$ contains all roots of $p(x)$ for all large enough a.

Obervation: for all large enough $a > 0$

$$\operatorname{sign}(p(a)) = \operatorname{sign}(\operatorname{lc}(p)),$$

$$\operatorname{sign}(p(-a)) = \begin{cases} \operatorname{sign}(\operatorname{lc}(p)), & \text{if } \deg(p) \text{ even;} \\ -\operatorname{sign}(\operatorname{lc}(p)), & \text{if } \deg(p) \text{ odd.} \end{cases}$$

Conclusion: for $S = \operatorname{Sturm}(p) = p_1, p_2, \ldots, p_n$

- $V_S(a)$ is variation in $\operatorname{lc}(p_1), \operatorname{lc}(p_2) \ldots, \operatorname{lc}(p_n)$; denoted $V_S(\infty)$.
- $V_S(-a)$ is variation in $\epsilon_1 \operatorname{lc}(p_1), \epsilon_2 \operatorname{lc}(p_2) \ldots, \epsilon_n \operatorname{lc}(p_n)$ where

$$\epsilon_i = \begin{cases} 1, & \text{if } \deg(p_i) \text{ even;} \\ -1, & \text{if } \deg(p_i) \text{ odd,} \end{cases}$$

denoted $V_S(-\infty)$.

Real Root Isolation by Continued Fractions

Method for isolating positive roots of $p(x)$.
Negative roots are positive roots of $p(-x)$.
Subdivide positive roots into: $(0,1)$, 1, $(1,\infty)$.
Express roots in $(0,1)$ as

$$\frac{1}{1+y}, \quad \text{for } y > 0.$$

To find all y clear denominator in $p(1/1+y)$ to get polynomial $p_I(y)$.
Express roots in $(1,\infty)$ as

$$1+y, \quad \text{for } y > 0.$$

To find all possible y put $p_T(y) = p(1+y)$.

Example: $p(x) = x^3 - 7x + 7$. By inspection 1 is not a root.

$$p_I(y) = 7y^3 + 14y^2 + 7y + 1.$$

Can't have any positive roots.

$$p_T(y) = y^3 + 3y^2 - 4y + 1.$$

By inspection 1 is not a root of $p_T(y)$.

$$p_{TI}(z) = z^3 - z^2 - 2z + 1,$$
$$p_{TT}(z) = z^3 + 6z^2 + 5z + 1.$$

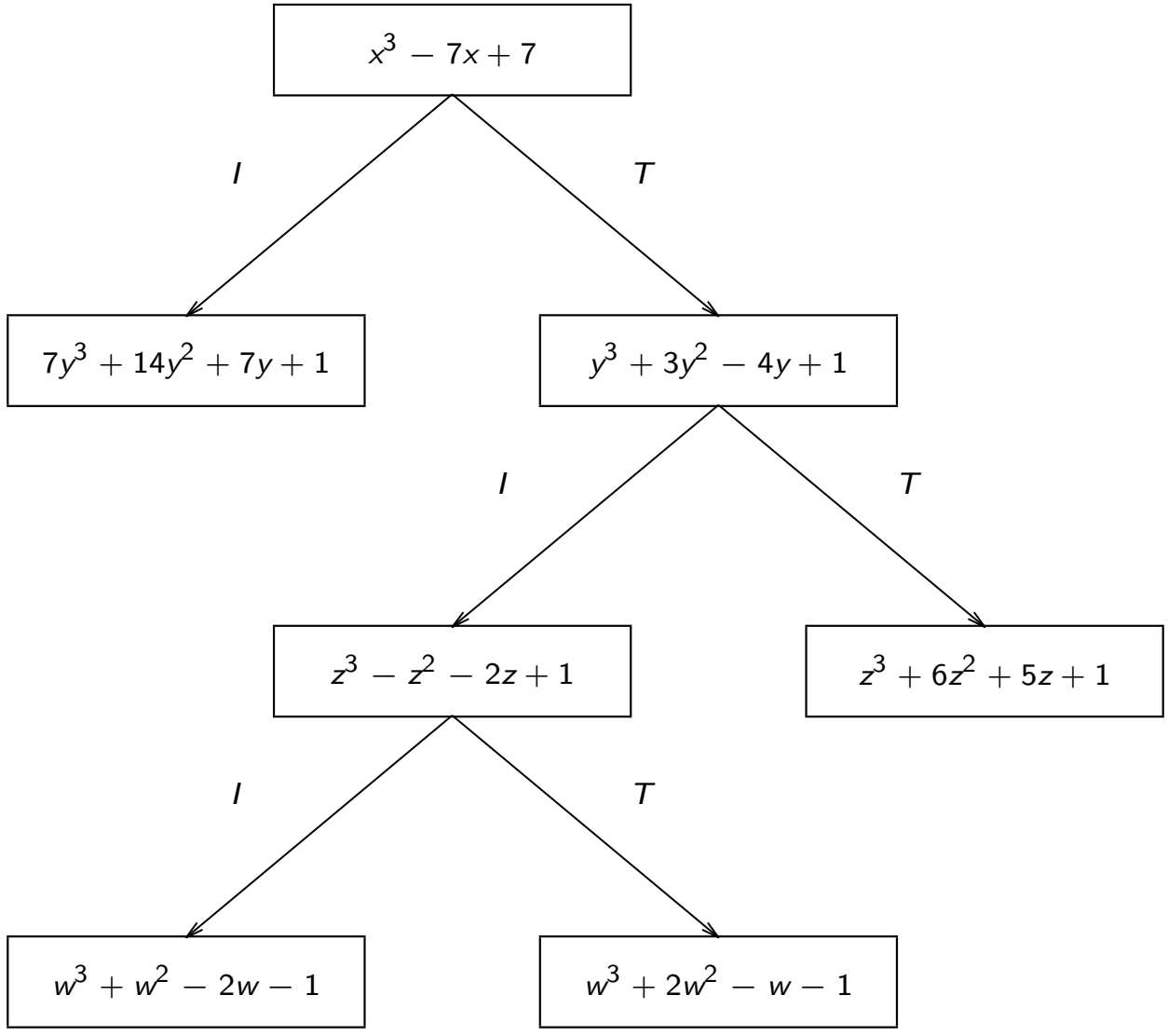
$p_{TT}(z)$ has no positive roots so $p_T(y)$ has no roots in $(1, \infty)$.
By inspection 1 is not a root of $p_{TI}(z)$. For roots in $(0, 1)$:

$$p_{TII}(w) = w^3 + w^2 - 2w - 1.$$

For roots in $(1, \infty)$:

$$p_{TIT} = w^3 + 2w^2 - w - 1.$$

Both have unique positive root.



Positive roots of $x^3 - 7x + 7$ isolated by $(1, 3/2)$ and $(3/2, 2)$.

Möbius Transforms

$$M(x) = \frac{a_1 x + a_0}{b_1 x + b_0}, \quad a_1 b_0 - a_0 b_1 \neq 0.$$

Transform $p(x)$ to

$$p_M(x) = p(M(x)) \cdot (b_1 x + b_0)^m, \quad m = \deg(p).$$

1. $p_M(\alpha) = 0 \Leftrightarrow p(M(\alpha)) = 0$ if $b_1\alpha + b_0 \neq 0$.
2. $a_1 b_0 - a_0 b_1 > 0 \Rightarrow M(x)$ strictly increasing.
3. $a_1 b_0 - a_0 b_1 < 0 \Rightarrow M(x)$ strictly decreasing.

Conclusion: If (a, b) isolates a root of $p_M(x)$ then interval with endpoints $M(a)$, $M(b)$ isolates corresponding root of $p(x)$.

Theorem: (Vincent, 1836) Let $p(x)$ be a square free polynomial with rational coefficients. Consider a sequence of transformations of $p(x)$ by

$$x \mapsto a_1 + \frac{1}{x}, \quad x \mapsto a_2 + \frac{1}{x}, \quad x \mapsto a_3 + \frac{1}{x}, \quad \ldots$$

where a_1 is an arbitrary non-negative integer and $a_2, a_3, \ldots$ are arbitrary positive integers. Then after finitely many steps the sequence of coefficients of the transformed polynomial has either zero or one sign variation.

Theorem: Let $p(x)$ be a polynomial with real coefficients that have exactly one sign variation. Then $p(x)$ has exactly one positive real root.

Speeding Things Up

Instead of

$$x \mapsto 1 + x$$

use

$$x \mapsto b + x$$

with b a good estimate of integer part of smallest positive root of $p(x)$.

Computing b: Use 2nd theorem of Cauchy to find upper bound B on positive roots of $x^m p(1/x)$. Take $b = \lfloor 1/B \rfloor$.

Theorem: (Budan,1807) Let $p(x)$ be square free of degree $m > 0$ and let a, b be two real numbers with $a < b$. Let V_a, V_b be the variation of the coefficients of $p(a+x)$, $p(b+x)$ respectively. Let r be the number of real roots of $p(x)$ in (a, b). Then

1. $V_a \geq V_b$.
2. $r \leq V_a - V_b$.
3. $(V_a - V_b) - r$ is an even number.

Corollary: If the coefficients of $q(x)$ and $q(1+x)$ have the same number of sign variations then $q(x)$ has no roots in $(0, 1)$.

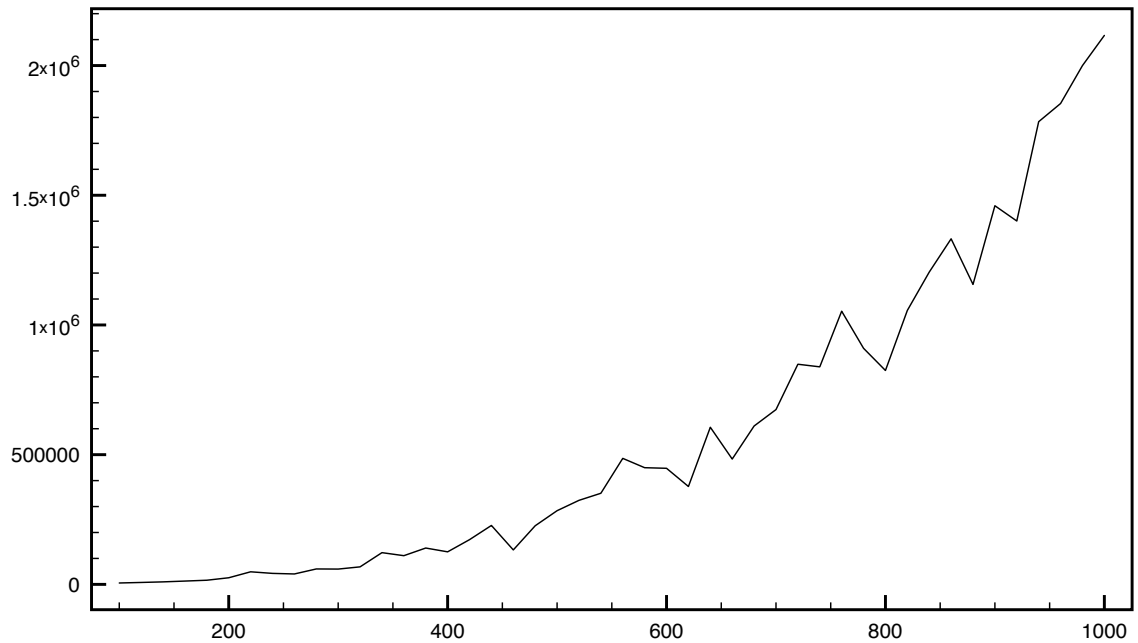
$CFISOL(p(x)) \mapsto [E, A]$

($p(x)$ is square free. E is a list of (some of the) exact non-negative roo
isolating intervals for the rest of the non-negative roots of $p(x)$.)

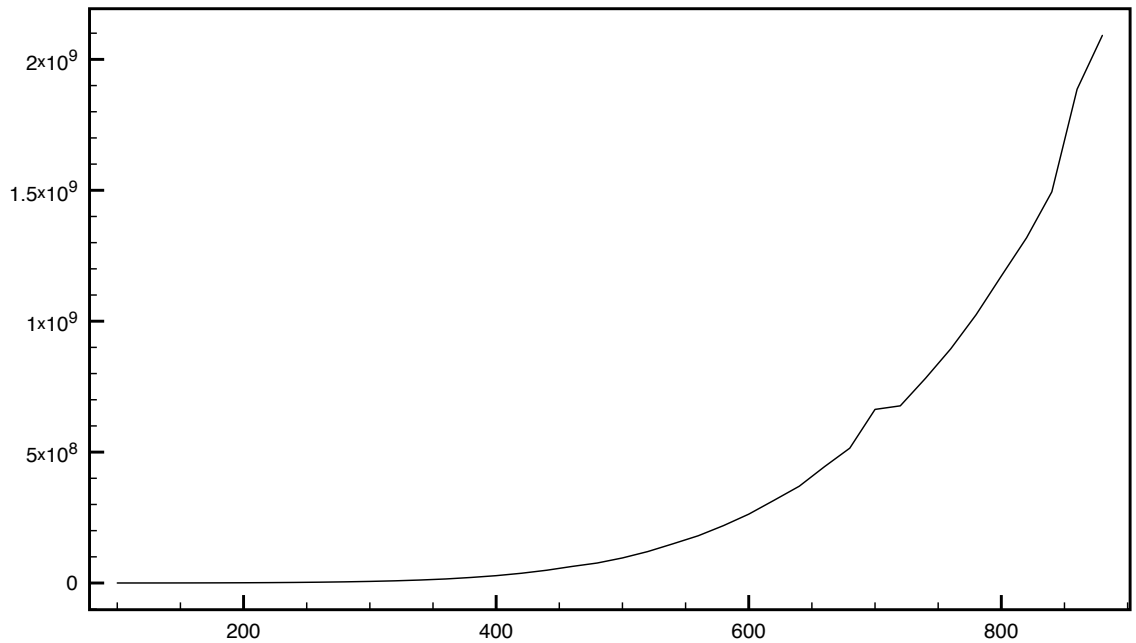
1. **if** $p(0) = 0$ **then**
$E := [0]$; $A := [\,]$; $p_w(x) := p(x)/x$
else $E := [\,]$; $A := [\,]$; $p_w(x) := p(x)$
fi;

2. $v := VAR(p_w(x))$;
if $v > 1$ **then** $T := [[p_w(x), x, v]]$ (work to be done)
elif $v = 1$ **then**
$T := [\,]$;
$u := UBPR(p_w(x))$;
if $p_w(u) = 0$ **then** $E := [op(E), u]$
else $A := [(0, u)]$
fi
fi;

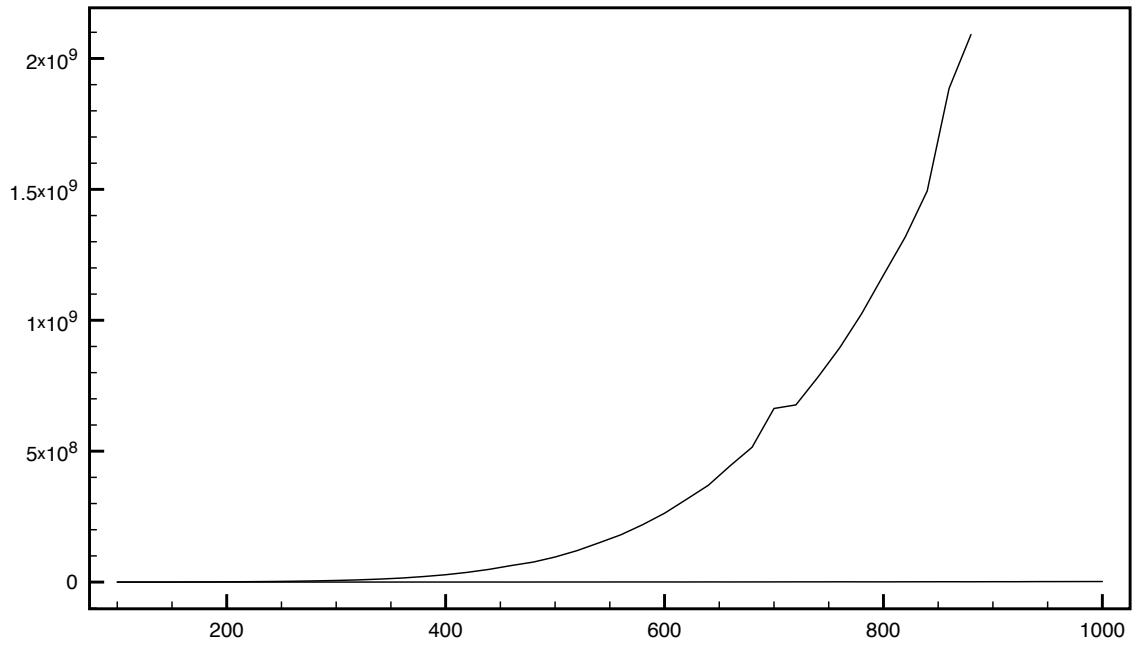
3. **while** $T \neq [\,]$ **do**
remove the first element $[p_M(x), M(x), v_M]$ from T;
$b := LBPR(p_M(x))$;
if $b \geq 1$ **then**
$p_M(x) := Moebius(p_M(x), b + x)$;
$M(x) := M(b + x)$;
if $p_M(0) = 0$ **then**
$E := [op(E), M(0)]$; $p_M(x) := p_M(x)/x$
fi;
fi;
$v_1 := v_M$;
$p_{M_1} := Moebius(p_M(x), 1 + x)$;
$M_1 := M(1 + x)$;
if $p_{M_1}(0) = 0$ **then**
$E := [op(E), M_1(0)]$; $p_{M_1}(x) := p_{M_1}(x)/x$
fi;
$v_{M_1} := VAR(p_{M_1})$;
if $v_{M_1} > 1$ **then** $T := [[p_{M_1}(x), M_1(x), v_{M_1}], op(T)]$
elif $v_{M_1} = 1$ **then**
$A := [op(A), make\_interval(p_{M_1}(x), M_1(x))]$
fi;
if $v_1 \neq v_{M_1}$ **then** ($p_M(x)$ might have roots in $(0, 1)$)
$p_I(x) := Moebius(p_M(x), 1/x)$;
if $\mathrm{lc}(p_I(x)) < 0$ **then** $p_I(x) := -p_I(x)$ **fi**;
$M_I := M(1/x)$;
$p_{M_1} := Moebius(p_I(x), 1 + x)$;
$M_1 := M_I(1 + x)$;
$v_{M_1} := VAR(p_{M_1}(x)$;
if $v_{M_1} > 1$ **then** $T := [[p_{M_1}(x), M_1(x), v_{M_1}], op(T)]$
elif $v_{M_1} = 1$ **then**
$A := [op(A), make\_interval(p_{M_1}(x), M_1(x))]$
fi
fi
od



Random polynomials with coefficients in $[-99, 99]$. Time in μs.



Collins-Krandick polynomials, $A_n = x^n - 2(x^2 - 3x + 1)^2$.



Random compared with Collins-Krandick polynomials.