The role of algorithms in symbolic computation

Erich Kaltofen
North Carolina State University
www.kaltofen.us

Caviness's foreword to the Computer Algebra Handbook

Two ideas lie gleaming on the jeweler's velvet. The first is the calculus, the second, the algorithm. The calculus and the rich body of mathematical analysis to which it gave rise made modern science possible; but it has been the algorithm that has made possible the modern world.

—David Berlinski, *The Advent of the Algorithm*

Caviness's foreword to the Computer Algebra Handbook

Two ideas lie gleaming on the jeweler's velvet. The first is the calculus, the second, the algorithm. The calculus and the rich body of mathematical analysis to which it gave rise made modern science possible; but it has been the algorithm that has made possible the modern world.

—David Berlinski, *The Advent of the Algorithm*

So, gentle reader, I recommend this volume and all its concepts, symbols, and algorithms to you.

—Bob Caviness, *Computer Algebra Handbook*

Important algorithms: “classical” computer algebra

Euclid, Chinese remainder

Sturm chains, Seidenberg’s algorithm

Gauss’s distinct degree factorization, Berlekamp/Zassenhaus

Berlekamp/Massey

Gröbner, Macaulay resultants, Wu triangular sets

Risch integration and transcendence theory of special functions

FFT-based polynomial arithmetic

Gosper and Karr

Collins cylindrical algebraic decomposition

. . .

Important algorithms: “middle earth”

Zippel and Ben-Or-Tiwari sparse interpolation

Singer and Kovacic differential equation solvers

Lattice basis reduction [LLL]

Zeilenberger

Wiedemann, block Wiedemann/Lanczos, matrix Padé

Straight-line and black box polynomial factorization

Baby steps/giant steps algorithms for lin. and polynomial algebra

Tellegen’s principle

Real roots of polynomial systems

Corless et al. approximate GCD, Sasaki approx. factorization

. . .

Important algorithms: “modern” symbolic computation

Sparse resultants, A- and J-resultants

Giesbrecht/Mulders-Storjohann diophantine linear solvers

Fast bit complexity in linear algebra over the integers

Black box matrix preconditioners, early termination

van Hoeij power sums, Bostan et al. logarithmic derivatives

Sparsest shift of polynomials

Villard-Jeannerod optimal polynomial matrix inverse

Skew, Ore and differential polynomial factorization

Approximate polynomial factorization via differential equations

Barvinok-Woods and De Loera et al. short rational functions

Lenstra lacunary polynomial factorization

…

Supersparse (lacunary) polynomials

The supersparse polynomial

$$f(X_1,\ldots,X_n) = \sum_{i=1}^{t} a_i X_1^{\alpha_{i,1}} \cdots X_n^{\alpha_{i,n}}$$

is input by a list of its coefficients and corresponding term degree vectors.

$$\text{size}(f) = \sum_{i=1}^{t} \Big(\text{size}(a_i) + \lceil \log_2(\alpha_{i,1} \cdots \alpha_{i,n} + 2) \rceil \Big)$$

Term degrees can be very high, e.g., $\geq 2^{500}$

Supersparse (lacunary) polynomials

The supersparse polynomial

$$f(X_1, \ldots, X_n) = \sum_{i=1}^{t} a_i X_1^{\alpha_{i,1}} \cdots X_n^{\alpha_{i,n}}$$

is input by a list of its coefficients and corresponding term degree vectors.

$$\mathrm{size}(f) = \sum_{i=1}^{t} \Big(\mathrm{size}(a_i) + \lceil \log_2(\alpha_{i,1} \cdots \alpha_{i,n} + 2) \rceil \Big)$$

Term degrees can be very high, e.g., $\geq 2^{500}$
Over $\mathbb{Z}_p$: evaluate by repeated squaring
Over $\mathbb{Q}$: cannot evaluate in polynomial-time exept for $X_i = 0, \pm 1$

Easy problems for supersparse polynomials $f = \sum_i a_i X^{\alpha_i} \in \mathbb{Z}[z]$

Cucker, Koiran, Smale 1998: Compute root $a \in \mathbb{Z}: \; f(a) = 0$.

Gap idea: if $f(a) = 0, a \neq \pm 1$ then $g_1(a) = \cdots = g_s(a) = 0$
where $f(X) = \sum_j g_j(X) X^{\alpha_j}$ and $\alpha_{j+1} - \alpha_j - \deg(g_j) \geq \chi$.

Easy problems for supersparse polynomials $f = \sum_i a_i X^{\alpha_i} \in \mathbb{Z}[z]$

Cucker, Koiran, Smale 1998: Compute root $a \in \mathbb{Z}: \ f(a) = 0$.

Gap idea: if $f(a) = 0, a \neq \pm 1$ then $g_1(a) = \cdots = g_s(a) = 0$
where $f(X) = \sum_j g_j(X) X^{\alpha_j}$ and $\alpha_{j+1} - \alpha_j - \deg(g_j) \geq \chi$.

Write $f(X) = \underbrace{g(X)}_{\deg(g) \leq k} + X^u h(X), \quad \|f\|_1 = |a_1| + \cdots + |a_t|$.

For $a \neq \pm 1, h(a) \neq 0$:

$$\begin{aligned} |g(a)| &< \|f\|_1 \cdot |a|^k \\ |a^u h(a)| &\geq |a|^u \end{aligned}$$

Easy problems for supersparse polynomials $f = \sum_i a_i X^{\alpha_i} \in \mathbb{Z}[z]$

Cucker, Koiran, Smale 1998: Compute root $a \in \mathbb{Z}: \;\; f(a) = 0$.

Gap idea: if $f(a) = 0, a \neq \pm 1$ then $g_1(a) = \cdots = g_s(a) = 0$
where $f(X) = \sum_j g_j(X) X^{\alpha_j}$ and $\alpha_{j+1} - \alpha_j - \deg(g_j) \geq \chi$.

Write $f(X) = \underbrace{g(X)}_{\deg(g) \leq k} + X^u h(X), \quad \|f\|_1 = |a_1| + \cdots + |a_t|$.

For $a \neq \pm 1, h(a) \neq 0$:

$$\begin{aligned} |g(a)| &< \|f\|_1 \cdot |a|^k \\ |a^u h(a)| &\geq |a|^u \end{aligned}$$

$$u - k \geq \chi = \log_2 \|f\|_1 \Longrightarrow |a|^u \geq 2^\chi \cdot |a|^k \geq \|f\|_1 \cdot |a|^k \Longrightarrow f(a) \neq 0.$$

Polynomial time root-finder uses the fact that for

$$g_j(X) = c_1 + c_2 x^{\beta_2} + \cdots + c_s x^{\beta_s}, \quad \beta_i - \beta_{i-1} < \chi, \quad s \le t$$

we have

$$\beta_i \le (i-1)(\chi-1),$$

so

$$\deg(g_j) \le (t-1)(\chi-1)$$

Generalization by H. W. Lenstra, Jr. 1999

Input: $\varphi(\zeta) \in \mathbb{Z}[\zeta]$ monic irred.; let $K = \mathbb{Q}[\zeta]/(\varphi(\zeta))$
a supersparse $f(X) = \sum_{i=1}^{t} a_i X^{\alpha_i} \in K[X]$
a factor degree bound d

Output: a list of all irreducible factors of f over K of degree $\leq d$
and their multiplicities (which is $\leq t$ except for X)

Let $D = d \cdot \deg(\varphi)$
There are at most $O(t^2 \cdot 2^D \cdot D \cdot \log(Dt))$ factors of degree $\leq d$

Bit complexity is $\big(\underline{t + \log(\deg f)} + \log \|f\| + \log \|\varphi\|\,\big)^{O(D)}$

Special case $\varphi = \zeta - 1, d = D = 1$: Algorithm finds all rational roots in polynomial-time.

Linear and quadratic bivariate factors

Input: a supersparse $f(X,Y) = \sum_{i=1}^{t} a_i X^{\alpha_i} Y^{\beta_i} \in \mathbb{Z}[X,Y]$ that is monic in X;
an error probability $\varepsilon = 1/2^l$

Output: a list of polynomials $g_j(X,Y)$ with $\deg_X(g_j) \leq 2$ and $\deg_Y(g_j) \leq 2$;
a list of corresponding multiplicities.

The g_j are with probability $\geq 1 - \varepsilon$ all irreducible factors of f over $\mathbb{Q}$ of degree ≤ 2 together with their true multiplicities.

Linear and quadratic bivariate factors

Input: a supersparse $f(X,Y) = \sum_{i=1}^{t} a_i X^{\alpha_i} Y^{\beta_i} \in \mathbb{Z}[X,Y]$ that is monic in X;
an error probability $\varepsilon = 1/2^l$

Output: a list of polynomials $g_j(X,Y)$ with $\deg_X(g_j) \leq 2$ and $\deg_Y(g_j) \leq 2$;
a list of corresponding multiplicities.

The g_j are with probability $\geq 1-\varepsilon$ all irreducible factors of f over $\mathbb{Q}$ of degree ≤ 2 together with their true multiplicities.

Bit complexity: $\left(t + \log(\deg f) + \log \|f\| + \log 1/\varepsilon\right)^{O(1)}$

Algorithm

Step 0: compute all factors of f that are in $\mathbb{Q}[Y]$ by Lenstra's method on the coefficients of X^{α_i}

Step 1: compute linear and quadratic factors in $\mathbb{Q}[X]$ of $f(X,0)$, $f(X,1)$ and $f(X,-1)$ by Lenstra's method

Step 2: interpolate all factor combinations;
Test if $g(X,Y)$ divides $f(X,Y)$ by

$0 \equiv f(X,a) \bmod (g(X,a),p)$ where $a \in \mathbb{Z}$, p prime are random

Leading coefficient problem

If the leading (trailing) coefficient in X does not vanish for $Y = 0, \pm 1$, then one can impose *a factor* of the leading (trailing) coefficient on g.

Cannot interpolate factors of $\sum_i (X^{2d_i} - 1)(Y^{2e_i} - 1) f_i(X, Y)$

But we can compute all factors $Y - aX - b$ of all supersparse polynomials deterministically

Generalized gap theorem

$Y - aX - b$ divides $f(X,Y) \iff 0 = \sum_{i=1}^{t} a_i X^{\alpha_i}(aX+b)^{\beta_i}$

Write $f(X,Y) = g(X,Y) + Y^u h(X,Y)$ with $\deg_Y(g) \leq k$.
If

$$u - k > \chi = 15.45 \cdot \log_2(t \cdot \mathrm{Height}\, f)$$

where

$$\mathrm{Height}\, f = \max_i |a_i| \quad \text{provided} \quad \mathrm{GCD}_i(a_i) = 1,$$

then for rational $(a,b) \neq (0,0), (\pm 1, 0), (0, \pm 1)$

$$f(X, aX+b) = 0 \Longrightarrow g(X, aX+b) = 0 \text{ and } h(X, aX+b) = 0.$$

Generalized gap theorem

$Y - aX - b$ divides $f(X,Y) \iff 0 = \sum_{i=1}^{t} a_i X^{\alpha_i}(aX+b)^{\beta_i}$

Write $f(X,Y) = g(X,Y) + Y^u h(X,Y)$ with $\deg_Y(g) \le k$.
If

$$u - k > \chi = 15.45 \cdot \log_2(t \cdot \mathrm{Height}\, f)$$

where

$$\mathrm{Height}\, f = \max_i |a_i| \quad \text{provided} \quad \mathrm{GCD}_i(a_i) = 1,$$

then for rational $(a,b) \neq (0,0), (\pm 1, 0), (0, \pm 1)$

$$f(X, aX+b) = 0 \Longrightarrow g(X, aX+b) = 0 \text{ and } h(X, aX+b) = 0.$$

Note: $1 \Big/ \log_2 \min_{d \geq 5} (1 + \cos \frac{2\pi}{5})^{\frac{\lfloor d/5 \rfloor}{d-1}} < 15.45$

Polynomial time algorithm for unknown a, b uses the facts

1. gap is independent of $a, b \in \mathbb{Q}$

2. first splits into segments $g_j(X, Y)$ with $\deg_Y(g_j) \le (t-1)(\chi-1)$

3. switches roles of X, Y in each g_j and splits into segments $g_{j,\ell}$ with $\deg_X(g_{j,\ell}) \le (t-1)(\chi-1)$

Hard problems for supersparse polynomials in $K[X,Y]$

Theorem
The set of all monic (in X) irreducible supersparse polynomials in $K[X,Y]$ is NP-hard for $K = \mathbb{Q}$ and $K = \mathbb{F}_q$ for all p and all sufficiently large $q = p^k$, via randomized reduction.

Corollary
Suppose we have a Monte Carlo polynomial-time irreducibility test for monic supersparse polynomials in $\mathbb{F}_{2^k}[x,y]$ (for sufficiently large k).
Then large integers can be factored in Las Vegas polynomial-time.

Supersparse integers

$2^{2^{2478782}}+1$ is divisible by $3\cdot 2^{2478785}+1$ [Cosgrave et al. 2003]

The factors of $2^{2^n}+1$ are primes of the form $k\cdot 2^{n+2}+1$
$641 = 5\cdot 2^7+1$ divides $2^{2^5}+1$ [Euler 1732]

Ron Rivest’s lessons learned (2002 Turing Award lecture)

- Try to solve “real-world” problems
- Moore’s law (#transistors/in^2 doubles every year) matters
- Theory matters
- Organizations matter

Moore's law and asymptotically fast algorithms

Strassen matrix multiplication
Knuth/Schönhage half GCD
baby-steps/giant-steps polynomial factorization
Tellegen's principle,. . .

are practical on today's problem sizes

Moore's law and asymptotically fast algorithms

Strassen matrix multiplication
Knuth/Schönhage half GCD
baby-steps/giant-steps polynomial factorization
Tellegen's principle,...

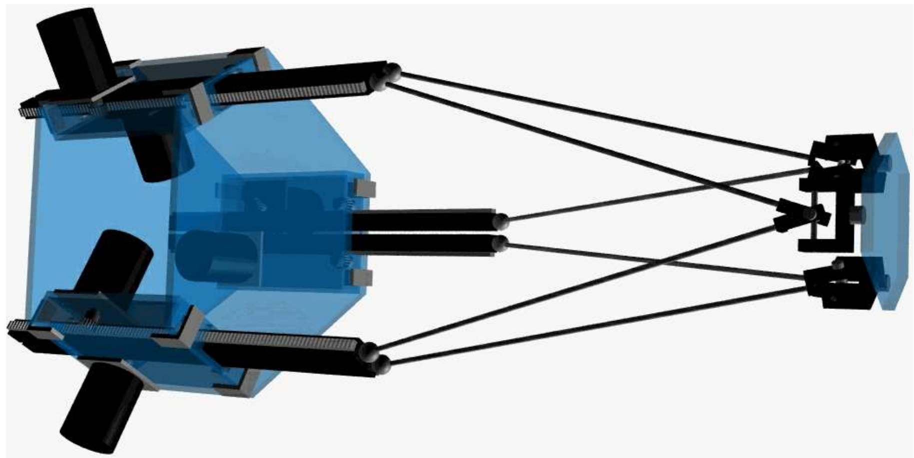
are practical on today's problem sizes

Abstract model vs. actual computer:
PRAM has $\Omega(\sqrt[3]{n})$ memory access time [D. Bernstein]
Turing algorithms have no cache faults [A. Schönhage]
$O(\log\log n)$ factors are model specific [E. Kaltofen]

Organizations

WRI and Maplesoft
ISSAC
JSC/AAECC
SIGSAM/Fachgruppe/JSSAC
W3C MathML committee

Real-world problems



Stewart platform: Josh Targownik's bypass surgery manipulator

Microwave antenna system: 4 equations in 4 variables of total degree 6 and 146 terms in total

The many medium size computations are real world applications