New Recombination Techniques for Polynomial Factorization Algorithms Based on Hensel Lifting

Grégoire Lecerf

CNRS UMR 8100
Laboratoire de Mathématiques
Université de Versailles St-Quentin-en-Yvelines
France
http://www.math.uvsq.fr/~lecerf

Context and Contents

Motivation. Design and implementation of efficient algorithms for factoring multivariate polynomials into irreducible factors.

Example. Irreducible factorization in

$$(\mathbb{F}_5(a,b,c)[d]/(d^2+a^2+2b^2-c^3))(e)[f]/(f^5+a^2+e^2+d)[y_1,y_2,y_3].$$

Recall that $F \in \mathbb{K}[y] \setminus \mathbb{K}$ is separable iff $\mathrm{Res}(F, F') \neq 0$, iff F has no multiple root in the algebraic closure $\bar{\mathbb{K}}$ of $\mathbb{K}$.

Contents.

I. Computability issues.

II. Reduction to separable polynomials.

III. Reduction from 2 to 1 variables.

Part I

Computability Issues

Effective field:= a field together with an implementation of its arithmetic operations ($+$, $-$, $\times$, $/$) and its equality test ($=$).

Negative Results

van der Waerden (1930), Fröhlich & Shepherdson (1956)

Theorem. The irreducible decomposition of univariate polynomials over effective fields is not computable in general.

Proof. Let $\lambda : \mathbb{N}^* \to \mathbb{N}^*$ be injective and computable. Let p_i be the ith prime number, and let $\mathbb{K} := \mathbb{Q}(\sqrt{p_{\lambda(1)}}, \sqrt{p_{\lambda(2)}}, \sqrt{p_{\lambda(3)}}, \ldots)$. For a given n, factoring $y^2 - p_n$ in $\mathbb{K}[y]$ is equivalent to testing if n is in the image of λ.

Take λ so that the latter test is not computable [Kleene, 1936].$\square$

Theorem. In characteristic $p > 0$, the pth power test, the pth root extraction and the squarefree decomposition are not computable.

Proof. $\mathbb{K} := \mathbb{F}_2(x_i, x_j^2 \mid i \notin \mathrm{Im}(\lambda),\ j \in \mathrm{Im}(\lambda)) \subseteq \mathbb{F}_2(x_1, x_2, \ldots)$.

x_n^2 is a square in $\mathbb{K}$ iff $n \notin \mathrm{Im}(\lambda)$. [von zur Gathen, 1984]

Remark. $\mathbb{K}$ is isomorphic to $\mathbb{F}_2(x_1, x_2, \ldots)$! $\square$

Positive Results

A field $\mathbb{K}$ is explicitly finitely generated over a field $\mathbb{F}$ if it is the fraction field of $\mathbb{F}[x_1, \ldots, x_n]/P$ with P *prime* and explicitly given by a finite set of generators.

Theorem. [van der Waerden, Fröhlich & Shepherdson, Seidenberg, Richman. . .] The irreducible decomposition is computable over any explicitly finitely generated extension of a prime field.

Proof. From now on, with a view towards complexity. . .

Theorem. [van der Waerden, Maclane, 30'] If $\mathbb{F}$ is perfect then, $\mathbb{K}$ can be rewritten into $\mathbb{K} = \mathbb{F}(t_1, \ldots, t_r)[\alpha_1, \ldots, \alpha_s]$, with

- $t_1, \ldots, t_r$ being a *transcendence basis* of $\mathbb{K}$ over $\mathbb{F}$,
- $\alpha_1, \ldots, \alpha_s$ being algebraic and *separable* over $\mathbb{F}(t_1, \ldots, t_r)$.

$\rightsquigarrow$ We can discard inseparable extensions.

Example. $\mathbb{K} := (\mathbb{F}_5(a, b, c)[d]/(d^2 + a^2 + 2b^2 - c^3))(e)[f]/(f^5 + a^2 + e^2 + d)$ can be rewritten into $\mathbb{F}_5(b, c, e, f)[d, a]/(d + f^5 + a^2 + e^2, a^2 + d^2 + 2b^2 - c^3)$.

Remarks.

☞ This rewriting can be made effective by means of Gröbner bases and pth root extractions in $\mathbb{F}_p$.

☞ After this rewriting, pth root extraction in $\mathbb{K}$ is made easier and boils down to linear algebra.

General Factorization Algorithms in Computer Algebra.

- Davenport, Trager (1981): never fully implemented and contained some gaps.
- Steel (2005): the first (and still unique) most general implementation in the `Magma` computer algebra system (`magma.maths.usyd.edu.au`).
- Other implementations are all partial.

Let us now move from computability to algorithms. . .

Prerequisites for Cost Analysis

[von zur Gathen and Gerhard, Modern Computer Algebra, 2003]

- Each binary arithmetic operation $(+, -, \times, /, =)$ in $\mathbb{K}$ costs $\mathcal{O}(1)$.
- Dense representation for polynomials.
  Example: the size of a bivariate polynomial of bi-degree (n, m) is $(n+1)(m+1)$.
- "Soft big Oh notation": $f(d) \in \tilde{\mathcal{O}}(g(d))$ means

$$f(d) \in g(d)(\log_2(3 + g(d)))^{\mathcal{O}(1)}.$$

- "Softly linear in d"= $\tilde{\mathcal{O}}(d)$; "Softly quadratic in d"= $\tilde{\mathcal{O}}(d^2)\ldots$
- The product, the division and the extended gcd of two univariate polynomials of degree d over $\mathbb{K}$ take $\tilde{\mathcal{O}}(d)$ operations in $\mathbb{K}$.
- ω is a constant such that the product of two $n \times n$ matrices over $\mathbb{K}$ takes $\mathcal{O}(n^\omega)$ arithmetic operations in $\mathbb{K}$. For convenience we assume that $2 < \omega \leq 3$.

Factorization in $\mathbb{F}_p[y]$(and $\mathbb{F}_{p^k}[y]$)

- Early ideas: Gauss (1797), Galois (1830), Arwins (1918).
- 1st alg.: Berlekamp (1970), Zassenhaus (1969), Cantor & Zassenhaus (1981).
- Alg. from 90's: von zur Gathen, Shoup, Niederreiter, Gao, Kaltofen. . .

Factorization in $\mathbb{Q}[y]$

- First algorithm due to Kronecker (1882): exponential cost.
- Hensel (1918) lifting algorithm (already known by Gauss): exponential cost. Popularized in computer algebra by Zassenhaus (1969).
- First polynomial time algorithm by Lenstra & Lenstra & Lovász (1982): compute a complex root with sufficiently high precision in order to deduce its minimal polynomial by means of LLL.
- First *practical* polynomial time algorithm: van Hoeij (2002); then improvements and implementation by Belabas & van Hoeij & Klüners & Steel (2004): compute an approximate p-adic decomposition, and recombine the factors with LLL.

Separable Algebraic Extension

☞ *Reduction to the separable case to be presented in the next part of the talk.*

Theorem. Factorization of separable polynomials in $\mathbb{K}(x)[y] \Longrightarrow$ factorization of separable polynomials in $\mathbb{K}[\alpha][y]$ whenever α is algebraic separable over $\mathbb{K}$.

Proof. Let $F \in \mathbb{K}[z, y]$, and let q be the minimal polynomial of α.

Irr. decomposition of $F(\alpha, y) \Longleftrightarrow$ prime decomposition of $(q(z), F(z, y))$

$\Longleftrightarrow$ Irr. decomposition of $\mathrm{Res}_z(q(z), F(z, y - xz)) \in \mathbb{K}[x][y]$. $\square$

☞ van der Waerden in *Moderne Algebra* (1930) in characteristic 0.

☞ Trager (1976): algorithmic point of view; probabilistic faster approach in characteristic 0 by taking a random value for x in $\mathbb{K}$.

☞ Steel (2005): complete implementation in positive characteristic in Magma.

☞ Bostan, Flajolet, Salvy, Schost (2006): speed-up for computing the resultant.

Transcendental Extension

Theorem. Factorization of separable polynomials in $\mathbb{K}[y] \Longrightarrow$ factorization of separable polynomials in $\mathbb{K}[x][y]$. *(Proof in Part III of the talk.)*

Let $F \in \mathbb{K}[x, y]$ of total degree d and bidegree d_x, d_y.

1st period: Exponential Time Algorithms

- The first algorithm goes back at least to Kronecker:
  - ☞ substitution $x \leftarrow y^{d_y+1}$, univariate factorization in degree $\mathcal{O}(d_x d_y)$;
  - ☞ exponential cost in the recombination step.
- The Hensel lifting and recombination approach was studied in [Musser, 1973, 1975], [Wang, Rothschild, 1975], [Wang, 1978], [von zur Gathen, 1984], [Bernardin, 1999 (`Maple` implementation)]...
  - ☞ Univariate factorization in degree d_y.
  - ☞ The exponential cost is again in the recombination step.
  - ☞ The cost is polynomial in average over finite fields [Gao, Lauder, 2000].
- Absolute factorization via elimination following Emmy Noether's ideas.

2nd period: First Polynomial Time Algorithms

- The first deterministic polynomial time algorithm for when $\mathbb{K} = \mathbb{Q}$ is due to Kaltofen (1982). Several authors then contributed during the 80's for various $\mathbb{K}$: Lenstra, Kannan, Lovász, Chistov, Grigoviev, von zur Gathen...
  - ☞ Derived from the LLL algorithm; essentially cubic time.

3rd period: Efficient Polynomial Time Algorithms

- **First recent breakthrough.** Shuhong Gao's reduction to linear algebra (2003) via de Rham's cohomology: $\tilde{\mathcal{O}}((d_x d_y)^2)$ (softly quadratic) in characteristic 0 or large enough. Derived from Ruppert's absolute irreducibility test (1986, 99).

- **Second recent breakthroughs.** The first polynomial time Hensel lifting and recombination algorithm is due to [Belabas, van Hoeij, Klüners, Steel, 2004]: $\tilde{\mathcal{O}}(d_x d_y^3)$.

- [Bostan, Lecerf, Salvy, Schost, Wiebelt, 2004]: improvement to $\tilde{\mathcal{O}}(d^3)$ in characteristic 0 or large enough.

- **Mixing of the breakthroughs.** [Lecerf, part III of the talk]: "$\tilde{\mathcal{O}}(d_x d_y^2)$".

Purely Inseparable Algebraic Extension

Theorem. Factorization of separable polynomials in $\mathbb{K}[y]$ $\not\Longrightarrow$ factorization of separable polynomials in $\mathbb{K}[\alpha][y]$ if α is purely inseparable over $\mathbb{K}$.

Solution already presented. Rewrite $\mathbb{K}[\alpha]$ as an extension of its prime field in order to remove purely inseparable extensions.

Other possible solution.

1. Let q be the minimal polynomial of α over $\mathbb{K}$. Wlog we can assume that $q(\alpha) = \alpha^p - a$, with $a \in \mathbb{K} \setminus \mathbb{K}^p$.

2. From a separable $F(y) \in \mathbb{K}[\alpha][y]$ compute $\tilde{F}(y^p) = F(y)^p$.

   $\tilde{F}$ is *separable* hence can be factored in $\mathbb{K}[y]$.

3. Let G be an irreducible factor of $\tilde{F}$. If $G(y^p)$ is a pth power H^p in $\mathbb{K}[\alpha][y]$ then H is an irreducible factor of F. Otherwise $G(y^p)$ is an irreducible factor of F.

☞ "$\not\Longrightarrow$" becomes "$\Longrightarrow$" if $\mathbb{K}$ satisfies Seidenberg's condition P.

Seidenberg's condition P on $\mathbb{K}$: pth power test and pth root extraction are possible in any purely inseparable extension of $\mathbb{K}$.

$\iff$ pth root test and extraction are possible in any finite algebraic extension of $\mathbb{K}$.

$\iff$ pth root test and extraction are possible in any explicitly finitely generated field extension of $\mathbb{K}$.

$\iff$ squarefree factorization is possible in $\mathbb{L}[y]$ for any finite algebraic extension $\mathbb{L}$ of $\mathbb{K}$ [Gianni, Trager, 1996].

Reference. Factorization in constructive mathematics: Mines, Richman, and Ruitenburg, *A course in constructive algebra*, Springer-Verlag, 1988.

Algebraically Closed Field

The absolute decomposition of $F \in \mathbb{K}[x, y]$ is its decomposition in $\bar{\mathbb{K}}[x, y]$, where $\bar{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$.

Example. $F := y^4 + (2x + 14)y^2 - 7x^2 + 6x + 47 =$
$(y^2 + (1 - 2\sqrt{2})x - 16\sqrt{2} + 7)(y^2 + (1 + 2\sqrt{2})x + 16\sqrt{2} + 7)$.

Usual Representation of the Absolutely Irreducible Factors.

Assume that F is separable when seen in $\mathbb{K}(x)[y]$. The absolutely irreducible factors of F, written $F_1, \ldots, F_r$, and are usually represented by $\{(q_1, \tilde{F}_1), \ldots, (q_s, \tilde{F}_s)\}$, such that:

- $q_i \in \mathbb{K}[z] \setminus \mathbb{K}$, monic, separable.
- $\tilde{F}_i \in \mathbb{K}[x, y, z]$, with $\deg_z(\tilde{F}_i) \leq \deg(q_i) - 1$.
- $\deg(\tilde{F}_i(x, y, \alpha))$ is independent of the root α of q_i.
- $\{F_1, \ldots, F_r\} = \cup_{i=1}^{s} \{\tilde{F}_i(x, y, \alpha) \mid q_i(\alpha) = 0\}$.
- Irredundancy: $\sum_{i=1}^{s} \deg(q_i) = r$.

Example 1. If $F \in \mathbb{K}[y]$ is squarefree then we can take $s := 1$, $q_1(z)$ as the monic part of $F(z)$ and $\tilde{F}_1(y, z) := y - z$.

Example 2. If $\mathbb{K} := \mathbb{Q}$ and $F := y^4 + (2x + 14)y^2 - 7x^2 + 6x + 47$ then we can take $s := 1$, $q_1(z) := z^2 - 2$, $\tilde{F}_1(x, y, z) := y^2 + (1 - 2z)x - 16z + 7$.

Theorem. [Noether, 1922] For all $\mathbb{K}$ the absolutely irreducible decomposition of any separable polynomial F can be computed by means of arithmetic operations in $\mathbb{K}$ alone.

☞ Computing an algebraic extension of $\mathbb{K}$ containing all the absolute factors of F is actually very expensive and useless in many applications.

Noether, 1922

Schmidt, 1976

Heintz, Sieveking, 1981

Trager, 1984

Dicrescenzo, Duval, 1984

Kaltofen, 1985: poly time

von zur Gathen, 1985

Ruppert, 1986

Dvornicich, Traverso, 1987

Bajaj, Canny, Garrity, Warren, 1989

Duval, 1990

Kaltofen, 1995: cubic time

Ragot, 1997

Ruppert, 1999

Cormier, Singer, Ulmer, Trager, 2002

Galligo, Rupprecht, 2002

Coreless, Galligo, *et al.*, 2002

Rupprecht, 2004

Bronstein, Trager, 2003

Gao, 2003: softly quadratic time

Sommese, Verschelde, Wampler, 2004

Chèze, Galligo, 2004

Chèze, Lecerf, 2005: sub-quadratic

Part II

Reduction to Separable Polynomials

Let $\mathbb{A}$ be a unique factorization domain.

Let $\boldsymbol{F} \in \mathbb{A}[\boldsymbol{y}]$ be primitive of degree $\boldsymbol{d}$.

$\boldsymbol{p}$ denotes the characteristic of $\mathbb{A}$.

$\boldsymbol{F} \in \mathbb{A}[\boldsymbol{y}] \setminus \mathbb{A}$ is said to be separable if it has no multiple root in the algebraic closure of the fraction field of $\mathbb{A} \Longleftrightarrow \mathbf{Res}(\boldsymbol{F}, \boldsymbol{F}') \neq \mathbf{0}$.

Definition

If $p = 0$ then separable decomposition $\equiv$ squarefree decomposition.

Now assume that $p > 0$.
The separable decomposition of F is the *unique* set
$\{(G_1, q_1, m_1), \ldots, (G_s, q_s, m_s)\} \subseteq (\mathbb{A}[y] \setminus \mathbb{A}) \times \{1, p, p^2, p^3, \ldots\} \times \mathbb{N}$
(the G_i are actually defined up to unit factors in $\mathbb{A}$) such that:

1. $F(y) = \prod_{i=1}^{s} G_i(y^{q_i})^{m_i}$;
2. for all $i \neq j$ in $\{1, \ldots, s\}$, $G_i(y^{q_i})$ and $G_j(y^{q_j})$ are coprime;
3. for all $i \in \{1, \ldots, s\}$, m_i mod $p \neq 0$;
4. for all $i \in \{1, \ldots, s\}$, G_i is separable and primitive;
5. for all $i \neq j$ in $\{1, \ldots, s\}$, $(q_i, m_i) \neq (q_j, m_j)$.

Proof. The roots of $G_i(y^{q_i})$ are the ones of F with multiplicity $q_i m_i$. $\square$

Algorithms

It is classical that the separable decomposition can be computed in polynomial time by arithmetic operations in $\mathbb{A}$ alone.

If $\mathbb{A}$ is a field:

- Gianni & Trager (1996): softly quadratic algorithm extending the classical squarefree factorization algorithm for characteristic 0 attributed to Musser (1971).

- Lecerf (2006): softly optimal cost, with a natural extension of Yun's squarefree factorization algorithm (1976).

Otherwise: the fast multimodular and Chinese remaindering techniques classically used for the $\gcd$ can be adapted to the separable factorization.

Reducing the Irreducible Factorization to the Separable Case

1. Compute the separable decomposition of F into $\prod_{i=1}^{s} G_i(y^{q_i})^{m_i}$.

2. Compute the irreducible factorization of each G_i.

3. If H is an irreducible factor of G_i then compute the largest $q|q_i$ such that $H(y^{q_i}) = P(y^{q_i/q})^q$. Then $P(y^{q_i/q})$ is an irreducible factor of F with multiplicity qm_i.

✎ pth power and pth root extraction must be computable.

Part III

Reduction from 2 to 1 Variables

Let $\boldsymbol{F} \in \mathbb{K}[\boldsymbol{x}, \boldsymbol{y}]$ be of total degree $\boldsymbol{d}$ and bi-degree $(\boldsymbol{d_x}, \boldsymbol{d_y})$.

$\boldsymbol{F}$ is assumed to be

- primitive when seen in $\mathbb{K}[\boldsymbol{x}][\boldsymbol{y}]$, and
- separable when seen in $\mathbb{K}(\boldsymbol{x})[\boldsymbol{y}]$.

The Classical Hensel Lifting Approach

Pretreatment

Task. Find a suitable translation of x so that the following normalization condition holds:

$$\deg_y(F(0,y)) = d_y \quad \text{and} \quad \operatorname{Res}_y\left(F, \frac{\partial F}{\partial y}\right)(0) \neq 0.$$

Algorithm. If $\mathbb{K}$ has sufficiently many elements then the translation can easily be found in $\mathbb{K}$ (softly optimal average cost). Otherwise we construct an algebraic extension $\mathbb{E}$ of $\mathbb{K}$ of degree $\tilde{\mathcal{O}}(\log(d_x d_y))$ in order to increase the cardinality. Then we compute the irreducible factorization of F in $\mathbb{E}[x,y]$ from which we deduce the one in $\mathbb{K}[x,y]$.

☞ The extra cost for working in $\mathbb{E}$ instead of $\mathbb{K}$ is negligible when discarding the logarithmic cost factors.

From now on we assume that the normalization condition holds.

Skeleton of the Hensel Lifting Factorization Algorithm

Let $F_1, \ldots, F_r$ be the irreducible factors of F.
Let c (resp. c_i) be the leading coefficient of F (resp. F_i) seen in $\mathbb{K}[x][y]$.
We write $F = c\mathfrak{F}_1 \cdots \mathfrak{F}_s$ for the irreducible factorization of F in $\mathbb{K}[[x]][y]$.
Each $\mathfrak{F}_i$ is made monic.

Algorithm.

1. Initialization: factor $F(0, y)$ in $\mathbb{K}[y]$ to obtain $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to precision (x).
2. Hensel lifting: use Hensel lifting in order to obtain $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to a certain precision (x^σ) (softly optimal cost).
3. Recombination: discover how the lifted factors recombine into the F_i.

Problem. Find an efficient polynomial time recombination.

For all $i \in \{1, \ldots, r\}$, let $\mu_i \in \{0, 1\}^s$ be the unique vector defined by

$$F_i = c_i \prod_{j=1}^{s} \mathfrak{F}_j^{\mu_{i,j}}.$$

☞ The knowledge of all the μ_i solves the recombination problem.

Example

$$F := y^4 - x^4 - 2y^3 + 2yx^2 - y^2 - x^2 + 2y \in \mathbb{Q}[x, y].$$

1. Initialization: $F(0, y) = y(y-1)(y+1)(y-2)$.
2. Hensel lifting:

$$\mathfrak{F}_1 = y - (2 - 1/2x^2 - 1/8x^4) + \mathcal{O}(x^5),$$

$$\mathfrak{F}_2 = y - (1 + 1/2x^2 - 1/8x^4) + \mathcal{O}(x^5),$$

$$\mathfrak{F}_3 = y - (1/2x^2 + 1/8x^4) + \mathcal{O}(x^5),$$

$$\mathfrak{F}_4 = y - (-1 - 1/2x^2 + 1/8x^4) + \mathcal{O}(x^5).$$

3. Recombination: $\mu_1 = (1, 0, 1, 0)$ and $\mu_2 = (0, 1, 0, 1)$.
   $F_1 = \mathfrak{F}_1\mathfrak{F}_3 = y^2 - 2y + x^2, \qquad F_2 := \mathfrak{F}_2\mathfrak{F}_4 = y^2 - x^2 - 1.$

$$F_i = c_i \prod_{j=1}^{s} \mathfrak{F}_j^{\mu_{i,j}}.$$

Detailed History of the Hensel Lifting Approach

Let σ still denote the precision of the lifted factors.

- Belabas, van Hoeij, Klüners, Steel (2004): logarithmic derivative method, $\sigma = d_x(2d_y - 1) + 1$ suffices to recombine in polynomial time.

  Theorem. $\mu_1, \ldots, \mu_r$ is the reduced echelon basis of the following system in the $\ell_i \in \mathbb{K}$: $\exists G \in \mathbb{K}[x, y]$, $\deg_x(G) \leq d_x$, $\deg_y(G) \leq d_y - 1$,

$$\sum_{i=1}^{s} \ell_i \frac{\frac{\partial \mathfrak{F}_i}{\partial y}}{\mathfrak{F}_i} - \frac{G}{F} \in (x^\sigma).$$

  ☞ The polynomial time was conjectured by T. Sasaki *et al.* (1991–1993) with a similar technique.

  ☞ The precision σ is sharp for this algorithm.

- Bostan, Lecerf, Salvy, Schost, Wiebelt (2004): $\sigma = 3d - 2$ suffices, if $\mathbb{K}$ has characteristic zero or at least $d(d-1) + 1$.

- Lecerf (2006): new algorithm based on the de Rham cohomology with precision $\sigma = 2d$, if $\mathbb{K}$ has characteristic zero or at least $d(d-1)+1$, and F monic in $\mathbb{K}[x][y]$.

  Theorem. $\mu_1, \ldots, \mu_r$ is the reduced echelon basis of the following system in the $\ell_i \in \mathbb{K}$: $\exists G, H \in \mathbb{K}[x, y],\ \deg(G) \leq d-1,\ \deg(H) \leq d-1,$

  $$\sum_{i=1}^{s} \ell_i \frac{\frac{\partial \mathfrak{F}_i}{\partial y}}{\mathfrak{F}_i} - \frac{G}{F} \in (x^\sigma) \text{ and } \sum_{i=1}^{s} \ell_i \frac{\frac{\partial \mathfrak{F}_i}{\partial x}}{\mathfrak{F}_i} - \frac{H}{F} \in (x^{\sigma-1}).$$

  ☞ The precision σ is also sharp for this algorithm.

- Lecerf (next slide): precision $\sigma = d_x + 1$ always suffices by means of a different recombination point of view.

The New Recombination Point of View

Let $\quad \hat{F}_i := \prod_{j=1, j\neq i}^{r} F_j = \dfrac{F}{F_i} \quad$ and $\quad \hat{\mathfrak{F}}_i := \dfrac{F}{\mathfrak{F}_i}.$

The central objects to recombine are the following:

$$\mathfrak{G}_i := \left\lceil \hat{\mathfrak{F}}_i \frac{\partial \mathfrak{F}_i}{\partial y} \right\rceil^{d_x+1} \text{ for all } i \in \{1, \ldots, s\},$$

where $\lceil A \rceil^l := \sum_{0 \leq i \leq l-1,\ j \geq 0} a_{i,j} x^i y^j$, for any $A := \sum_{i,j \geq 0} a_{i,j} x^i y^j$.

☞ The only lifting to precision (x^{d_x+1}) is necessary to compute the $\mathfrak{G}_i$.

Let $\mathbb{F}$ be a sub-field of $\mathbb{K}$.

$$\mathcal{L}_{\mathbb{F}} := \left\{ (\ell_1, \ldots, \ell_s) \in \mathbb{F}^s \mid \sum_{i=1}^{s} \ell_i \mathfrak{G}_i \in \left\langle \hat{F}_1 \frac{\partial F_1}{\partial y}, \ldots, \hat{F}_r \frac{\partial F_r}{\partial y} \right\rangle_{\mathbb{F}} \right\},$$

Lemma. $\mu_1, \ldots, \mu_r$ is the reduced echelon basis of $\mathcal{L}_{\mathbb{F}}$.

Proof. $F_i = c_i \prod_{j=1}^{s} \mathfrak{F}_j^{\mu_{i,j}} \Longrightarrow \hat{F}_i \dfrac{\partial F_i}{\partial y} = \sum_{j=1}^{s} \mu_{i,j} \hat{\mathfrak{F}}_j \dfrac{\partial \mathfrak{F}_j}{\partial y} \Longrightarrow \hat{F}_i \dfrac{\partial F_i}{\partial y} = \sum_{j=1}^{s} \mu_{i,j} \mathfrak{G}_j,$

whence $\mu_i \in \mathcal{L}_{\mathbb{F}}$. Then conclude with the dimensions... $\square$

Characterization of $\mathcal{L}_{\mathbb{F}}$ by the Residues

$\ell := (\ell_1, \dots, \ell_s) \in \mathbb{F}^s$, $\quad G := \sum_{i=1}^{s} \ell_i \mathfrak{G}_i$, $\quad \bar{\mathbb{K}} :=$ algebraic closure of $\mathbb{K}$.

Let $\phi_1, \dots, \phi_{d_y}$ be the roots of F in $\bar{\mathbb{K}}[[x]]$, and let $\rho_i := G(x, \phi_i) / \frac{\partial F}{\partial y}(x, \phi_i)$, for all $i \in \{1, \dots, d_y\}$, so that

$$\frac{G}{F} = \sum_{i=1}^{d_y} \frac{\rho_i}{y - \phi_i}.$$

Lemma. $\ell \in \mathcal{L}_{\mathbb{F}} \implies \rho \in \mathbb{F}^{d_y}$. Conversely, $\rho \in \bar{\mathbb{K}}^{d_y} \implies \ell \in \mathcal{L}_{\mathbb{F}}$.

Proof. If $\ell \in \mathcal{L}_{\mathbb{F}}$ then G is a $\mathbb{F}$ linear combination of $\hat{F}_1 \frac{\partial F_1}{\partial y}, \dots, \hat{F}_r \frac{\partial F_r}{\partial y}$. Conversely . . . $\square$

We shall distinguish two cases:

a. Characteristic $p = 0$ or $p \geq d_x(2d_y - 1) + 1$.

☞ We take $\mathbb{F} = \mathbb{K}$; $\qquad \rho \in \mathbb{K}^{d_y} \iff d(\rho)/dx = 0$.

b. $0 < p \leq d_x(2d_y - 1)$.

☞ We take $\mathbb{F} = \mathbb{F}_p$; $\qquad \rho \in \mathbb{F}_p^{d_y} \iff (\rho_i)^p = \rho_i$ for all i.

Computation of $\mathcal{L}_{\mathbb{K}}$ in characteristic 0

$$\mathbf{D}: \quad \mathbb{K}[x,y]_{d_x,d_y-1} \longrightarrow \mathbb{K}(x)[y]_{d_y-1}$$

$$G \mapsto \left(\frac{\partial G}{\partial x}\frac{\partial F}{\partial y} - \frac{\partial G}{\partial y}\frac{\partial F}{\partial x}\right)\frac{\partial F}{\partial y} - \left(\frac{\partial^2 F}{\partial xy}\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial y^2}\frac{\partial F}{\partial x}\right) G \operatorname{mod}_y F,$$

$$\frac{d\rho_i}{dx} = \frac{\mathbf{D}(G)(x,\phi_i(x))}{\frac{\partial F}{\partial y}(x,\phi_i(x))^3},$$

Proposition. $\langle \mu_1, \ldots, \mu_r \rangle = \mathcal{L}_{\mathbb{K}} = \ker(\mathbf{D})$.

Warning. F is not monic when seen in $\mathbb{K}[x][y]$. In order to avoid expression swell "$\operatorname{mod}_y F$" is performed in $\mathbb{K}[[x]][y]$ to precision $(x^{\mathcal{O}(d_x)})$. In this way the recombination reduces to the resolution of a linear system with s unknowns and $\mathcal{O}(d_x d_y)$ equations.

Deterministic Recombination Algorithm in characteristic 0

Input. $F \in \mathbb{K}[x, y]$, and $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to precision (x^{d_x+1}).

Output. $\mu_1, \ldots, \mu_r$.

1. For each $i \in \{1, \ldots, s\}$, compute $\hat{\mathfrak{F}}_i$ as the quotient of F by $\mathfrak{F}_i$ to precision (x^{d_x+1}). $\tilde{\mathcal{O}}(sd_xd_y)$.

2. Compute $\hat{\mathfrak{F}}_1 \frac{\partial \mathfrak{F}_1}{\partial y}, \ldots, \hat{\mathfrak{F}}_s \frac{\partial \mathfrak{F}_s}{\partial y}$ to precision (x^{d_x+1}) and deduce $\mathfrak{G}_1, \ldots, \mathfrak{G}_s$. $\tilde{\mathcal{O}}(sd_xd_y)$

3. Compute $\mathbf{D}(\mathfrak{G}_1), \ldots, \mathbf{D}(\mathfrak{G}_s)$. $\tilde{\mathcal{O}}(sd_xd_y)$

4. Compute $\mu_1, \ldots, \mu_t$ as the reduced echelon solution basis of the following linear system in the unknowns $(\ell_1, \ldots, \ell_s) \in \mathbb{K}^s$:

$$\sum_{i=1}^{s} \ell_i \mathbf{D}(\mathfrak{G}_i) = 0. \qquad \tilde{\mathcal{O}}(d_xd_ys^{\omega-1})$$

The worst case for this deterministic algorithm is when $s \approx d_y \rightsquigarrow \tilde{\mathcal{O}}(d_x d_y^{\omega})$

If necessary, we can swap x and y in order to ensure $d_y \leq d_x$ so that $\tilde{\mathcal{O}}(d_x d_y^{\omega}) \subset \tilde{\mathcal{O}}((d_x d_y)^2) \rightsquigarrow$ softly quadratic cost.

First speedup:

The linear system to be solved is overdetermined: at most d_y unknowns for $\mathcal{O}(d_x d_y)$ equations.

$\rightsquigarrow$ Use a Las Vegas probabilistic linear solver [Kaltofen, Saunders, 1991] in order to reach an average total cost in $\tilde{\mathcal{O}}(d_x d_y^2) \subseteq \tilde{\mathcal{O}}((d_x d_y)^{1.5})$ (when $d_y \leq d_x$).

More speedups: many tricks can be used in order to make the cost of the linear algebra negligible in practice [Belabas et al., 2004], [Lecerf, 2005].

Computation of $\mathcal{L}_{\mathbb{F}_p}$ in characteristic $p > 0$

From now on we assume that $0 < p \leq d_x(2d_y - 1)$.
$\mathbb{K}[x, y]_{k,l} :=$ polynomials of bi-degree at most (k, l).

We use the Niederreiter (1993) operator:

$$\tilde{\mathbf{N}}: \quad \mathbb{K}[x, y]_{d_x, d_y - 1} \longrightarrow \mathbb{K}[x, y^p]_{p d_x, d_y - 1}$$

$$G \mapsto G^p + \frac{\partial^{p-1}}{\partial y^{p-1}} \left(F^{p-1} G\right).$$

WARNING: $\tilde{\mathbf{N}}$ is not $\mathbb{K}$-linear in general but only $\mathbb{F}_p$-linear.

$$\mathbf{N}: \quad \mathbb{F}_p^s \longrightarrow \mathbb{K}[x, y^p]$$

$$(\ell_1, \ldots, \ell_s) \mapsto \mathbf{N}\left(\sum_{i=1}^{s} \ell_i \mathfrak{G}_i\right).$$

Proposition. $\mu_1, \ldots, \mu_r$ is the reduced echelon basis of $\ker(\mathbf{N})$.
Proof. The same as for polynomials in $\mathbb{F}_p[y]$... $\square$

☞ The recombination problem reduces to linear system solving over $\mathbb{F}_p$.

☞ The size of the linear system to be solved depends on the $\mathbb{F}_p$-algebra structure of $\mathbb{K}$.

☞ If $\mathbb{K} = \mathbb{F}_{p^k}$ then the linear system has $\mathcal{O}(pkd_xd_y)$ equations and s unknowns.

Proposition.

- $\ker(\mathbf{N}) \subseteq \ker(\mathbf{D}) \cap \mathbb{F}_p^s$.
- $\mathbf{N}(\ker(\mathbf{D}) \cap \mathbb{F}_p^s) \subseteq \mathbb{K}[x^p, y^p]_{d_x, d_y - 1}$.

☞ If $\mathbb{K} = \mathbb{F}_{p^k}$ then the new linear system has $\mathcal{O}(kd_xd_y)$ equations and s unknowns.

Sketch of the Recombination Algorithm

1. Run the algorithm designed for the characteristic 0 in order to get a basis of $\ker(\mathbf{D}) \cap \mathbb{F}_p^s$.

2. Compute the reduced echelon basis of $\ker(\mathbf{N})$.

When $\mathbb{K} = \mathbb{F}_q$ with $q = p^k$ we have the following estimates:

- Deterministic version: $\tilde{\mathcal{O}}(k d_x d_y^\omega)$ operations in $\mathbb{F}_p$ "$\leq$" $\tilde{\mathcal{O}}(d_x d_y^\omega)$ operations in $\mathbb{F}_q$.

- Randomized version: average cost in $\tilde{\mathcal{O}}(k d_x d_y^2)$ operations in $\mathbb{F}_p$ "$\leq$" $\tilde{\mathcal{O}}(d_x d_y^2)$ operations in $\mathbb{F}_q$.

Conclusion

Future work.

- Extension of [Chèze, Lecerf, 2005]: absolute factorization in small positive characteristic, and a unified approach of the rational and the absolute factorizations.
- Generalization of the complexity results in terms of the volume of the convex hull of the support of F.
- Implementation of an open source `C++/Mathemagix` factorization library (`www.mathemagix.org` [van der Hoeven]).