Journal of Symbolic Computation 44 (2009) 1194–1211

Contents lists available at ScienceDirect

Journal of Symbolic Computation

journal homepage: www.elsevier.com/locate/jsc

Towards toric absolute factorization[☆]

M. Elkadi, A. Galligo, M. Weimann

Université de Nice Sophia Antipolis, Laboratoire J-A. Dieudonné, Parc Valrose, 06108 Nice Cedex2, France

ARTICLE INFO

Article history:
Received 13 November 2007
Accepted 28 March 2008
Available online 26 February 2009

Keywords:
Absolute polynomial factorization
Irreducible factors
Bivariate
Algebraic interpolation
Toric geometry
Newton polytope
Semi-numerical probabilistic algorithm

ABSTRACT

This article presents an algorithmic approach to study and compute the absolute factorization of a bivariate polynomial, taking into account the geometry of its monomials. It is based on algebraic criterions inherited from algebraic interpolation and toric geometry.

1. Introduction

Multivariate polynomial factorization and production of software dedicated to the effective solving of this problem has received much attention in Computer Algebra. Whereas rational factorization is only concerned with factors in $\mathbb{Q}[\underline{x}] := \mathbb{Q}[x_1, \ldots, x_n]$, absolute factorization provides all irreducible factors with coefficients in $\overline{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$. The bivariate case contains most of the difficulties of the problem. In theory, by Bertini's theorem and via Hensel liftings, the multivariate problem reduces to the bivariate one. In the present article we will concentrate on the bivariate case but our techniques naturally extend to the n variables' case, $n \geq 2$.

Polynomial absolute factorization has been considered from many points of view, see e.g. Bostan et al. (2004), Chèze and Galligo (2005) and Chèze and Lecerf (2007) and their bibliography, but during the last decade two main strategies have been quite successful. On the one hand, an algebraic approach relies on the study of Ruppert–Gao matrix (Ruppert, 1986; Gao, 2003). It has been improved in

[☆] This research was partly supported by the French ANR, contract Gecko. The two first authors are also members of the Galaad project team, INRIA-UNS.

E-mail addresses: elkadi@math.unice.fr (M. Elkadi), galligo@math.unice.fr (A. Galligo), weimann@ujf-grenoble.fr (M. Weimann).

0747-7171/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsc.2008.03.007

Chèze and Lecerf (2007) and Lecerf (2007) to provide an algorithm with a quasi-optimal complexity. On the other hand, a geometric approach, based on a zero-sum criterion, provides very efficient semi-numerical probabilistic algorithms able to deal with polynomials having degree up to 200, see Rupprecht (2004), Chèze (2004) and Chèze and Galligo (2005). This criterion was derived from the study of the monodromy group, of a projection of the curve C defined by f (the polynomial to be factorized) on a line, acting on a smooth fiber. A similar strategy was developed and implemented in Sommese et al. (2001, 2004), and its use was extended for obtaining the irreducible decomposition of an algebraic set. The zero-sums considered in Sasaki et al. (1991) admit more general interpretations in Algebraic Geometry as traces.

The model of computation in these approaches is the following. The input is a polynomial with integer coefficients and the output is a list of polynomials with coefficients in an algebraic extension of $\mathbb{Q}$ which should also be computed. In order to determine these coefficients the strategy consists in embedding $\overline{\mathbb{Q}}$ in $\mathbb{C}$ and representing approximations of these coefficients by bigfloats. Then conjugacy relations are used to recognize an algebraic presentation of an extension of $\mathbb{Q}$ and an exact algebraic presentation of the coefficients. This model of computation is commonly used in Number Theory and was successfully adapted in Rupprecht (2004) and improved in Chèze and Galligo (2006) by considering a representation with algebraic integers. In this paper we focus on the geometric foundation needed for our generalization together with some illustrative examples. So we will not provide precise bounds for our approximations nor details for efficient implementations. These tasks will be addressed in a future work.

The aim of this article is first to reinterpret the vanishing traces criterions in the geometric approach as a consequence of Wood's theorem (Wood, 1984) on algebraic interpolation of a family of analytic germs of curves. Second, to provide a generalization of Wood's theorem inspired by Weimann (submitted for publication) and adapted to the factorization of polynomials with fixed Newton polytopes. Third, to outline an algorithm for toric absolute factorization that we tested on examples.

When a polynomial f of total degree d is given by the collection of its coefficients which are all nonzero, its representation is called dense. Whereas when some coefficients of f are known to be zero, its Newton polytope (i.e. the convex hull of exponents of monomials of its nonzero coefficients) is considered and its representation is called toric or sparse. Then adapted algorithms are developed; e.g. toric elimination received much attention (Gelfand et al., 1994; Emiris, 1996).

To the best of our knowledge most of the existing articles on polynomial factorization deal with dense polynomials, although in Abu Salem et al. (2004) a study of toric rational polynomial factorization was presented; it is based on adapted Hensel liftings. Our aim is to rely on this article: assuming that f is already irreducible in $\mathbb{Q}[\underline{x}]$, we compute its absolute factorization. In that case, all Newton polytopes of absolute factors of f are equal and are homothetic to that of f. Hence the combinatorial task is simplified and the difficulty concentrates on the geometry with a fixed toric variety. Let us also mention von zur Gathen and Kaltofen (1985) where the multivariate sparse factorization is reduced to the dense bivariate or univariate polynomial factorization.

The paper is organized as follows. In the next section, the special shape of absolute factors of an irreducible rational polynomial is shown. Section 3 explains the use of interpolation of analytic germs of curves via a Burger's PDE to derive a vanishing trace criterion in $\mathbb{P}^2(\mathbb{C})$, and compares it with the use of a monodromy action. Section 4 generalizes this trace criterion to a (possibly singular) toric surface. Section 5 outlines an algorithm for toric absolute factorization. It is based on an algebraic criterion inherited from interpolation problems in toric geometry, and computations of traces. It generalizes and improves the algorithm developed for dense polynomials in Rupprecht (2004) and Chèze and Galligo (2005); its different steps are illustrated on an example. The algorithm for the case of bidegree polynomials is presented, for these polynomials it improves significantly the one in Rupprecht (2004) and Chèze and Galligo (2005) . Finally, remarks and hints for future improvements are listed in a conclusion. At the end of this paper a short Appendix collects the properties of abstract toric surfaces needed for our developments.

Hereafter $\mathbb{P}^n$ denotes the projective space over $\mathbb{C}$ of dimension n. For a polynomial map (f, q) in $\mathbb{C}^2$, $\mathrm{Jac}(f, q)$ is its Jacobian. The Newton polytope of a polynomial f is denoted by N_f and the mixed volume of two polytopes P and Q by $\mathrm{MV}(P, Q)$.

2. Factorization and Newton polytopes

The Minkowski sum of two polytopes P and Q is

$$P + Q = \{p + q : p \in P, q \in Q\}.$$

The following two results are needed for our developments.

Proposition 1 (*Ostrowksi Theorem (Ostrowski, 1975)*). *The Newton polytope of the product of two polynomials g and h is the Minkowski sum of Newton polytopes of its factors:* $N_{gh} = N_g + N_h$.

So if the irreducible polynomial $f \in \mathbb{Q}[\underline{x}]$ has a polytope which is integrally indecomposable, f is absolutely irreducible. Let us also mention a study (Gao, 2001) of the irreducibility of a polynomial from a Newton polytope point of view.

Proposition 2. *Let* $f \in \mathbb{Q}[\underline{x}]$ *be an irreducible polynomial and* $f = f_1 \dots f_q$ *be its absolute factorization. Then the irreducible absolute factors* f_i *of* f *are conjugate over* $\mathbb{Q}$.

Proof. Up to a linear change of coordinates, f can be assumed monic in x_2, and consequently its absolute factors are also monic in x_2. Let G be the Galois group of the smallest extension of $\mathbb{Q}$ containing all the coefficients of f_1. If $\sigma \in G$, the conjugate polynomial $\sigma(f_1)$ of f_1 also divides f. Now as f is an irreducible element in $\mathbb{Q}[\underline{x}]$, the polynomial $\prod_{\sigma \in G} \sigma(f_1) = f$, hence each absolute factor f_j of f is equal to $\sigma(f_1)$ for some $\sigma \in G$. □

The determination of Newton polytopes of absolute factors of an irreducible polynomial in $\mathbb{Q}[\underline{x}]$ is highly simplified by the following corollary.

Corollary 3. *Let* $f \in \mathbb{Q}[\underline{x}]$ *be an irreducible polynomial and* $f = f_1 \dots f_q$ *be its absolute factorization. Then* $N_{f_1} = \dots = N_{f_q}$ *and* $N_f = qN_{f_1}$.

So, a polynomial $f \in \mathbb{Q}[x]$ of bidegree (d_1, d_2) which is irreducible over $\mathbb{Q}$ is irreducible over $\mathbb{C}$ if d_1 and d_2 are relatively prime.

Remark 4. Proposition 2 implies that the absolute factorization of f is completely determined by the number of factors q, an irreducible univariate polynomial $g(t) \in \mathbb{Q}[t]$ defining a finite extension $\mathbb{K} = \mathbb{Q}[t]/\big(g(t)\big)$, and the coefficients of f_1 which belong to $\mathbb{K}$ and are indexed by the lattice points in the polytope $\frac{1}{q}N_f \subset \mathbb{N}^2$.

3. Factorization and algebraic interpolation

Let f be an irreducible bivariate rational polynomial of total degree $d \geq 2$. This property implies that f is reduced over $\mathbb{C}$, then its absolute irreducible factors are in one-to-one correspondence with irreducible components of the affine curve C defined by f:

$$C = \{(x_1, x_2) \in \mathbb{C}^2 : f(x_1, x_2) = 0\}.$$

Sard–Bertini theorem combined with Bézout's theorem ensures that for $t = [t_0 : t_1 : t_2]$ generic in the dual projective space $(\mathbb{P}^2)^*$, the affine line

$$L_t = \{(x_1, x_2) \in \mathbb{C}^2 : t_0 + t_1 x_1 + t_2 x_2 = 0\}$$

intersects C transversely in d distinct points whose coordinates vary holomorphically with t by the implicit function theorem. Thus L_t defines a degree d reduced 0-cycle of C:

$$L_t \cdot C = p_1(t) + \dots + p_d(t).$$

The principle of uniqueness of analytic continuation and Bézout's theorem imply that f admits a factor of degree $k \leq d$ if and only if there exists

$$I = \{i_1, \dots, i_k\} \subset \{1, \dots, d\}$$

and an algebraic curve $C_I \subset \mathbb{C}^2$ of degree k such that (as shown in Fig. 1) for t in a small open set of $(\mathbb{P}^2)^*$:

$$L_t \cdot C_I = p_{i_1}(t) + \dots + p_{i_k}(t).$$

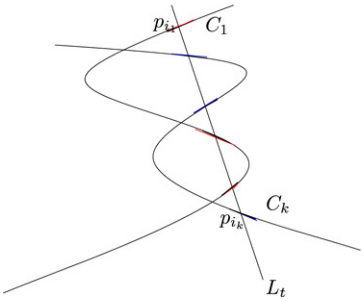

Fig. 1.

This is closely related to the (classical) problem: let $t \in (\mathbb{P}^2)^*$ distinct from the point at infinity $[1 : 0 : 0]$ and let $C_1 \cup \cdots \cup C_k$ be an union of germs of smooth analytic curves (algebraic in our case) of $\mathbb{C}^2$ transverse to the line L_t at pairwise distinct points $p_{i_1}(t), \ldots, p_{i_k}(t)$. Does there exist an algebraic curve of total degree k which contains all these germs C_i?

The following result solves that problem.

Theorem 5 (*Wood's Theorem (Wood, 1984)*). *The union of analytic curves $C_1 \cup \cdots \cup C_k$ is contained in an algebraic curve of degree k if and only if the germ of holomorphic function trace on the first coordinate, defined by*

$$(Tr\, x_1)(t) := \sum_{j=1}^{k} x_1(p_{i_j}(t))$$

is affine in the constant coefficient t_0 of L_t.

Geometrically, this result asserts that an analytic curve is algebraic if and only if the barycenters of intersection points with a generic line L lie on a line (called a diameter of the curve, see the line D in Fig. 2) when L moves parallel to itself, as shown in Fig. 2. Newton had already remarked this property in Newton (1710) for algebraic plane curves of degree 3. The proof of Theorem 5 in Wood (1984) is simple but relies on a tricky use of a Burger's PDE. It will be generalized for our purpose in Section 4.

In Rupprecht (2004) and Chèze and Galligo (2005) an algorithm for absolute dense factorization was developed based on vanishing partial sums. This algorithm uses topological considerations about the complex plane $\mathbb{C}^2$. Its proof relies on Harris uniform position theorem and Van Kampen theorem which establish the link between the irreducibility of an affine algebraic curve and the transitive action of a monodromy group (see Chèze and Galligo (2005) for details). It turns out that this condition on vanishing partial sums is equivalent to the interpolation criterion given by Wood's theorem. Let us recall briefly the principle of this method. Up to a linear change of variables, we assume that f is monic as a polynomial in x_2 of degree d. For $x_1 = a$ generic, let $x_{2,1}(a), \ldots, x_{2,d}(a)$ be the roots of the univariate polynomial $f(a, x_2)$. For each $i = 1 \ldots d$, let

$$\phi_i(x_1) = \sum_j \alpha_{j,i}(a)(x_1 - a)^j$$

be the power series satisfying $\phi_i(a) = x_{2,i}(a)$ and $f\big(x_1, \phi_i(x_1)\big) = 0$. Then $f(x) = f(x_1, x_2) = \prod_{i=1}^{d}\big(x_2 - \phi_i(x_1)\big)$. Every absolute factor of f has the form

$$f_I = \prod_{i\in I}\big(x_2 - \phi_i(x_1)\big) = x_2^{\delta} + a_{I,1}(x_1)x_2^{\delta-1} + \cdots + a_{I,\delta}(x_1),$$

with $I \subset \{1, \ldots, d\}$, $\mathrm{card}(I) = \delta$ and $\deg a_{I,i}(x_1) \le i$ for $i = 1 \ldots \delta$. In particular, the degree of $a_{I,1}(x) = -\sum_{i\in I}\phi_i(x_1)$ is at most 1, then $\sum_{i\in I}\alpha_{2,i}(a) = 0$. Because of the genericity, it turns out that this last condition is also sufficient for f to have an absolute factor. So in order to find absolute factorization of f it suffices to search minimal zero-sums between the complex numbers $\alpha_{2,1}(a), \ldots, \alpha_{2,d}(a)$.

The brute force resulting algorithm requires 2^d trace tests to detect factors of f. Strategies relying on LLL were developed and implemented in Chèze (2004) to decrease this high number of tests.

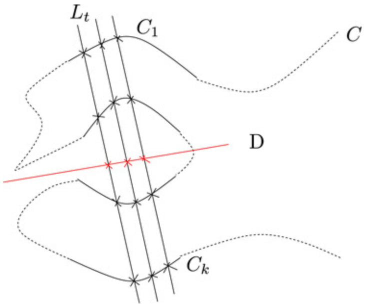

Fig. 2.

4. Interpolation in toric surfaces

In Weimann (submitted for publication) a necessary and sufficient condition was given for a family of germs of analytic hypersurfaces in a smooth projective toric variety X to be interpolated by an algebraic hypersurface with a prescribed class in the Chow ring of X. Here we establish a similar result in a toric surface which can be singular. This generalization is needed for our factorization algorithm.

4.1. Toric surfaces

Let us denote by $\mathbb{T}$ the algebraic torus $(\mathbb{C}^*)^2$. The Newton polytope P of a Laurent polynomial f gives information about the asymptotic behavior of the curve

$$C := \{x \in \mathbb{T}, f(x) = 0\}.$$

We say that a curve $D \subset \mathbb{T}$ is supported by an integer convex polytope Q if it is the zero set of a Laurent polynomial with Newton polytope Q.

Let Q be an integer convex polytope such that $Q \cap \mathbb{Z}^2 = \{m_0, \ldots, m_l\}$. Consider the morphism

$$\begin{aligned} \phi_Q &: \mathbb{T} \longrightarrow \mathbb{P}^l \\ x = (x_1, x_2) &\longmapsto [x^{m_0} : \cdots : x^{m_l}]. \end{aligned}$$

The Zariski closure X_Q of $\phi_Q(\mathbb{T}) \subset (\mathbb{C}^*)^l$ in $\mathbb{P}^l$ is the projective toric variety associated to Q. See Fulton (1993) or the Appendix at the end of this paper where the definition of an abstract toric surface and some of its properties are provided.

Without loss of generality we assume that $m_0 = 0$. We have $\dim X_Q = \dim Q$.

Lemma 6. *If* $\dim Q = 2$, *the map* ϕ_Q *is an embedding.*

D. Cox indicated (without proof) at the end of the survey paper (Cox, 2003) that this result is known. For a proof, we refer to Corollary 1.3.4. in Bruns et al. (1997).

4.2. Traces for curves in toric surfaces

4.2.1. Notations

Here we set notations that we will follow in the paper.

Let $Q \subset \mathbb{R}^2$ be a two-dimensional integer convex polytope with lattice points $m_0 = 0, m_1, \ldots, m_l$. Let $X = X_Q$ be the projective toric surface associated to Q. As seen in the previous section, ϕ_Q is one-to-one. Let $[u_0 : \cdots : u_l]$ be homogeneous coordinates on $\mathbb{P}^l$.

Every Laurent polynomial

$$q_a(x) = \sum_{i=0}^{l} a_i x^{m_i}$$

supported by Q determines a curve $C_a := \{q_a = 0\} \subset \mathbb{T}$. Since ϕ_Q is one-to-one, by Lemma 13 in Appendix, for a generic, C_a can be identified with the hyperplane section of $X \cap (\mathbb{C}^*)^l$ defined by the projective hyperplane

$$H_a = \{u \in \mathbb{P}^l : \sum_{i=0}^{l} a_i u_i = 0\}.$$

We denote by $a = [a_0 : \cdots : a_l]$ the point of the dual space $(\mathbb{P}^l)^*$ corresponding to C_a. For the definition of the mixed volume in the following lemma and its properties, see Gelfand et al. (1994).

Lemma 7. *Let $C \subset \mathbb{T}$ be a reduced curve supported by a lattice polytope P. For $a \in (\mathbb{P}^l)^*$ generic, C_a is smooth, irreducible and intersects C transversely in $d = \mathrm{MV}(P, Q)$ distinct points $p_1(a), \ldots, p_d(a)$, where $\mathrm{MV}(P, Q)$ denotes the mixed volume of (P, Q).*

Proof. Let us denote by $\mathcal{C}$ and $\mathcal{C}_a$ the Zariski closure in X of the affine curves $\phi_Q(C)$ and $\phi_Q(C_a)$. We know from Lemma 13 in Appendix that $\mathcal{C}_a$ coincides for generic a with the hyperplane section $H_a \cap X$ of X. Thus Bertini's theorem implies that the curve $\mathcal{C}_a$ is generically smooth irreducible and intersects $\mathcal{C}$ in its Zariski open set $\phi_Q(C)$. Since by Lemma 6 ϕ_Q is an embedding, we deduce that C_a is generically smooth, irreducible and intersects C transversely in $d = \deg(H_a \cdot X \cdot \mathcal{C})$ points. Bernstein's theorem asserts that $d = \deg\big(\mathcal{O}_X(1)\big)_{|\mathcal{C}} = \mathrm{MV}(P, Q)$. □

From this lemma, we have the following definition.

Definition 8. For any holomorphic function h near $C_\alpha \cap C$, the trace of h on C relative to the polytope Q is

$$(\mathrm{Tr}_C h)(a) := \sum_{j=1}^{d} h(p_j(a)).$$

This function is defined and holomorphic for a near α.

4.2.2. A necessary condition to interpolate germs of curves

We provide a necessary condition for a family of germs of curves to be interpolated by an algebraic curve C.

Since $m_0 = 0$ is a vertex of the polytope Q, the generic polynomial q_a has a nonzero constant term a_0.

Theorem 9. *Let $C \subset \mathbb{T}$ be an algebraic curve, and $\alpha \in (\mathbb{P}^l)^*$ satisfying the hypothesis of Lemma 7. We denote by Γ the union of facets of Q not containing 0. For $n \in \mathbb{N}^*$ and $s \in n(Q \cap \mathbb{Z}^2)$, we have*

$$\partial_{a_0}^{(n)}\big(\mathrm{Tr}_C\, x^s\big) = 0 \quad \text{if } s \in n(Q \setminus \Gamma), \tag{1}$$
$$\partial_{a_0}^{(n+1)}\big(\mathrm{Tr}_C\, x^s\big) = 0 \quad \text{if } s \in n\Gamma.$$

Proof. Suppose that $C = \{f = 0\}$, for a Laurent polynomial $f = \sum c_m x^m$. The trace function $\mathrm{Tr}_C\, x^s$ is a rational function on $\mathbb{P}^l$, it is homogeneous of degree 0 in a. If we denote by res_p and Res respectively the local Grothendieck residues at p and the global Grothendieck residue (see Griffiths and Harris (1978), Section 5), then for a in a small neighborhood of α,

$$\big(\mathrm{Tr}_C\, x^s\big)(a) = \sum_{p \in \mathbb{T}} \mathrm{res}_p \frac{x^s \mathrm{d}f \wedge \mathrm{d}q_a}{f q_a} = \mathrm{Res}\begin{bmatrix} x^s df \wedge dq_a \\ f \quad q_a \end{bmatrix}. \tag{2}$$

Since

$$\mathrm{d}f \wedge \mathrm{d}q_a = \left(\sum_{(m,m_i)} a_i\, c_m \det(m, m_i) x^{m+m_i-(1,1)}\right) \mathrm{d}x_1 \wedge \mathrm{d}x_2,$$

we obtain

$$\left(\mathrm{Tr}_C\, x^s\right)(a) = \sum_{(m,m_i)} a_i\, c_m \det(m, m_i)\ \mathrm{Res}\begin{bmatrix} x^{s+m+m_i}\,\frac{dx_1\wedge dx_2}{x_1x_2} \\ f \quad q_a \end{bmatrix}. \tag{3}$$

Using Cauchy formula for residues and Stokes theorem (Griffiths and Harris, 1978),

$$\partial_{a_0}^{(n)}\left(\mathrm{Res}\begin{bmatrix} x^{s+m+m_i}\,\frac{dx_1\wedge dx_2}{x_1x_2} \\ f \quad q_a \end{bmatrix}\right) = (-1)^n\, n!\ \mathrm{Res}\begin{bmatrix} x^{s+m+m_i}\,\frac{dx_1\wedge dx_2}{x_1x_2} \\ f \quad q_a^{n+1} \end{bmatrix}. \tag{4}$$

If P^0 denotes the interior of a polytope P, then by the toric version of Abel–Jacobi theorem (Hovanskiĭ, 1978), we have

$$s+m+m_i \in \left(N_f + (n+1)Q\right)^0 \implies \mathrm{Res}\begin{bmatrix} x^{s+m+m_i}\,\frac{dx_1\wedge dx_2}{x_1x_2} \\ f \quad q_a^{n+1} \end{bmatrix} = 0, \tag{5}$$

where N_f is the Newton polytope of f.

Let us denote by $Q_1 = [0, s_1]$ and $Q_2 = [0, s_2]$ the two facets of Q containing the origin 0, so that $Q = Q^0 \cup Q_1 \cup Q_2 \cup \Gamma$. To finish the proof we consider different cases:

(1) If $s \in (nQ)^0$, then for all $m \in N_f$ and $m_i \in Q$,

$$s+m+m_i \in \left(N_f + (n+1)Q\right)^0, \quad \text{so that } \partial_{a_0}^{(n)}(\mathrm{Tr}_C\, x^s) = 0.$$

(2) Let $s \in Q_1 \setminus \{ns_1\} = [0, ns_1[$. Since we are dealing with residues in the torus, we check easily that (2) depends on f up to multiplication by any Laurent monomial. Thus we can assume that N_f is contained in the cone generated by Q and intersects the ray $\mathbb{R}^+s_1$ in a nonempty set $N \subset N_f$ (consisting in one vertex or one facet of N_f). In this case, it is easy to check that for all $m \in N_f$ and $m_i \in Q$ such that $m+m_i \notin \mathbb{R}^+s_1$, $s+m+m_i \in \left(N_f + (n+1)Q\right)^0$. Moreover, $m+m_i \in \mathbb{R}^+s_1$ if and only if m and m_i are in $\mathbb{R}^+s_1$, that is $\det(m, m_i) = 0$. The formulas (3) and (5) show that $\partial_{a_0}^{(n)}(\mathrm{Tr}_C\, x^s) = 0$.

The same argument holds for $s \in [0, ns_2[$.

(3) If $s \in n\Gamma \setminus \{ns_1, ns_2\}$, N_f is contained in the cone $\mathbb{R}^+Q$ and we check that for all $m_i \in Q$ and $m \in N_f$, $s+m+m_i \in \left(N_f + (n+2)Q\right)^0$, so $\partial_{a_0}^{(n+1)}(\mathrm{Tr}_C\, x^s) = 0$.

(4) Let $s = ns_1$, as for the case 2, we choose $N_f \subset \mathbb{R}^+Q$ such that $N = N_f \cap \mathbb{R}^+s_1$ is nonempty. Then we check easily that if m or m_i does not belong to $\mathbb{R}^+s_1$, $s+m+m_i \in \left(N_f + (n+2)Q\right)^0$ and if m and m_i are in $\mathbb{R}^+s_1$, $\det(m, m_i) = 0$. So when $s = ns_1$, $\partial_{a_0}^{(n+1)}(\mathrm{Tr}_C\, x^s) = 0$.

The same argument is valid for $s = ns_2$.

These items combined with (3)–(5) imply (1). □

4.3. Criterion for algebraic interpolation

Now we give a necessary and sufficient criterion of interpolation generalizing Theorem 5 to our setting. To simplify the exposition and without loss of generality we further assume that the vectors $m_1 \in Q$ and $m_2 \in Q$ generate the lattice $\mathbb{Z}^2$ and $a_1, \ldots, a_t$ code the vertices of Q other than 0. Hence $a_{t+1}, \ldots, a_l$ code the other points of Q, where $l = \mathrm{card}(Q \cap \mathbb{Z}^2) - 1$.

Theorem 10. *Let $\alpha \in (\mathbb{P}^l)^*$ such that $C_a \subset \mathbb{T}$ is an irreducible smooth curve supported by Q for any a near α. Let*

$$C = C_1 \cup \cdots \cup C_d$$

be a union of germs of smooth analytic curves at pairwise distinct points $p_1, \ldots, p_d$ of C_α. Suppose that none of the germs C_i is contained in a curve $\{x^{m_1} - c = 0\}$, $c \in \mathbb{C}^*$. Then, there exists an algebraic curve

$\widetilde{C} \subset \mathbb{T}$, containing C and supported by a polytope P whose mixed volume with Q is d, if and only if, for generic $(a_1, \ldots, a_l)$ in a neighborhood of $(\alpha_1, \ldots, \alpha_l)$, the germ of holomorphic function

$$a_0 \longmapsto \left(\mathrm{Tr}_C\, x^{m_1}\right)(a_0)$$

is polynomial of degree at most 1 in the constant coefficient a_0.

Proof. Suppose that $\widetilde{C} = \{f = 0\}$, where f is a Laurent polynomial with Newton polytope P such that $\mathrm{MV}(P, Q) = d$. As $C \subset \widetilde{C}$, the two sets $C \cap C_a$ and $\widetilde{C} \cap C_a$ coincide for a in a sufficiently small neighborhood $U_\alpha \subset (\mathbb{P}^l)^*$ of α, since by Lemma 7 they have the same cardinal $d = \mathrm{MV}(P, Q)$. Thus, for $a \in U_\alpha$,

$$\forall s \in \mathbb{Z}^2, \quad \mathrm{Tr}_C\, x^s = \mathrm{Tr}_{\widetilde{C}}\, x^s,$$

and the necessary condition follows from Theorem 2.

Conversely, since the curve C_α is supported by Q, none of the coefficients $(\alpha_0, \ldots, \alpha_t)$ vanish.

Let us denote by $p_j(a)$ the intersection point of the germ C_j at α with C_a and we define the following germs of the holomorphic function at $\alpha \in (\mathbb{P}^l)^*$

$$X_i^{(j)}(a) := x^{m_i}(p_j(a)), \quad i = 0 \ldots l, j = 1 \ldots d.$$

We have

$$y \in C_j \cap C_a \Longrightarrow X_i^{(j)}\left(-\sum_{i=1}^{l} a_i y^{m_i}, a_1, \ldots, a_l\right) = y^{m_i}, \quad \forall a \in U_\alpha, \tag{6}$$

where U_α is a neighborhood of α. Differentiating the right-hand side of this implication according to a_1, we obtain:

$$\left(\partial_{a_1} X_i^{(j)} - y^{m_1} \partial_{a_0} X_i^{(j)}\right)\left(-\sum_{i=1}^{l} a_i y^{m_i}, a_1, \ldots, a_l\right) = 0.$$

Replacing $y \in C_j$ by $p_j(a) \in C_j$, and using the equality $-\sum_{i=1}^{l} a_i y^{m_i}\left(p_j(a)\right) = a_0$, we obtain a Burger's PDE:

$$\partial_{a_1} X_i^{(j)}(a) - X_1^{(j)}(a)\partial_{a_0} X_i^{(j)}(a) = 0.$$

So for $i = 1$,

$$\partial_{a_1} X_1^{(j)} = \frac{1}{2}\partial_{a_0}\left[X_1^{(j)}\right]^2.$$

This PDE is summable on j and gives rise to

$$\partial_{a_1}\left(\mathrm{Tr}_C\, x^{m_1}\right) = \frac{1}{2}\partial_{a_0}\left(\mathrm{Tr}_C\, x^{2m_1}\right).$$

We have a propagation of the behavior in the variable a_0: if $\mathrm{Tr}_C\, x^{m_1}$ is affine in a_0 then obviously $\partial_{a_1}(\mathrm{Tr}_C\, x^{m_1})$ is affine in a_0. By this PDE, $\partial_{a_0}(\mathrm{Tr}_C\, x^{2m_1})$ is also affine in a_0, hence the degree of $\mathrm{Tr}_C\, x^{2m_1}$ in a_0 equals at most 2. By induction on n, the map

$$a_0 \mapsto \mathrm{Tr}_C\, x^{nm_1}$$

is a polynomial of degree at most n in a_0.

Consider the following polynomial in X:

$$\begin{aligned} P(X, a) &:= \left(X - X_1^{(1)}(a)\right) \times \cdots \times \left(X - X_1^{(d)}(a)\right) \\ &= X^d - \sigma_1(a)X^{d-1} + \cdots + (-1)^d \sigma_d(a), \end{aligned}$$

the σ_i's are the elementary symmetric functions of $x^{m_1}\big(p_1(a)\big), \ldots, x^{m_1}\big(p_d(a)\big)$. Replacing a_0 by $-\sum_{i=1}^{l} a_i x^{m_i}$, and denoting $a' := (a_1, a_2, \ldots, a_l)$ and $a'' = (a_1, \ldots, a_t)$, we obtain a function

$$Q_{a'}(x) = \left(x^{m_1} - X_1^{(1)}\left(-\sum_{i=1}^{l} a_i x^{m_i}, a'\right)\right) \times \cdots \times \left(x^{m_1} - X_1^{(d)}\left(-\sum_{i=1}^{l} a_i x^{m_i}, a'\right)\right)$$

which vanishes on C for any a' near α', using (6).

Now, Newton formulas relating the coefficients of P with the traces of the power of the Laurent monomial x^{m_1} imply that the analytic functions

$$(a_0, a'') \mapsto \sigma_i(a_0, a'', \alpha_{t+1}, \ldots, \alpha_l)$$

are polynomial in a_0 (with degree at most n) for any $a'' := (a_1, \ldots, a_t)$ near α''. Thus the function

$$R_{a''}(x) := Q_{a'', \alpha_{t+1}, \ldots, \alpha_l}(x)$$

is a Laurent polynomial in x vanishing on C. So that the algebraic set defined by the following infinite number of equations:

$$\widetilde{C} := \{x \in \mathbb{T} : R_{a''}(x) = 0, \forall\, a'' \text{ near } \alpha''\}$$

contains C. We need to show that $C_a \cap \widetilde{C} = \{p_1(a), \ldots, p_d(a)\}$ for all a in a neighborhood of α. By construction, a point q belongs to $\widetilde{C} \cap C_\alpha$ if and only if there exists $j \in \{1, \ldots, d\}$ such that for all a'' near α''

$$x^{m_1}(q) = x^{m_1}\left(p_j\left(-a_1 x^{m_1}(q) - a_2 x^{m_2}(q) - \sum_{i=3}^{l} \alpha_i x^{m_i}(q), a_1, a_2, \alpha_3, \ldots, \alpha_l\right)\right). \tag{7}$$

Let us suppose that C_j is locally parameterized by

$$C_j = \{p(t), |t| < \epsilon,\ p(0) = p_j\}.$$

We consider the affine system in (a_0, a_2):

$$\begin{cases} a_0 + \alpha_1 x^{m_1}\big(p(t)\big) + a_2 x^{m_2}\big(p(t)\big) = c_p \\ a_0 + \alpha_1 x^{m_1}(q) + a_2 x^{m_2}(q) = c_q \end{cases} \tag{8}$$

where we define $c_q := -\sum_{i=3}^{l} \alpha_i x^{m_i}(q)$ for any $q \in \mathbb{T}$. Suppose that there exists $q \in C_\alpha \setminus \{p_j\}$ which satisfies (7). Then $x^{m_1}(q) = x^{m_1}(p_j)$ and, since m_1 and m_2 generate $\mathbb{Z}^2$, $q \neq p_j$ implies $x^{m_2}(q) \neq x^{m_2}(p_j)$. Thus, it is easy to check that there is a unique solution $(a_0(t), a_2(t))$ to (8) which converges to (α_0, α_2) when $|t|$ goes to zero. Thus, the map

$$a_2 \longmapsto p_j\big(-\alpha_1 x^{m_1}(q) - a_2 x^{m_2} - \sum_{i=3}^{l} \alpha_i x^{m_i}(q), \alpha_1, a_2, \alpha_3, \ldots, \alpha_l\big)$$

is surjective from a neighborhood of α_2 to C_j, so that

$$x^{m_1}(q) = x^{m_1}(p), \quad \forall\, p \in C_j.$$

This situation has been excluded by hypothesis. Thus we have proved that

$$\widetilde{C} \cap C_\alpha = C \cap C_\alpha.$$

By hypothesis, the last argument is valid when replacing $(\alpha_0, \ldots, \alpha_t)$ by a vector in its neighborhood. Thus for (a_0, a'') close to (α_0, α'')

$$\widetilde{C} \cap C_{a_0, a'', \alpha_{t+1}, \ldots, \alpha_l} = C \cap C_{a_0, a'', \alpha_{t+1}, \ldots, \alpha_l}.$$

Since the coefficients $(a_0, a_3, \ldots, a_t)$ correspond to the vertices of Q, the Zariski closure of $C_{a_0, a'', \alpha_{t+1}, \ldots, \alpha_l}$ in the toric variety $X = X_Q$ can avoid any finite subset of the divisor at infinity $X \setminus \mathbb{T}$ by choosing a generic value of (a_0, a''). Thus the Zariski closure in X of the two curves $\widetilde{C}$ and $C_{a_0, a'', \alpha_{t+1}, \ldots, \alpha_l}$ intersect transversely in the torus for a'' generic. This open condition remains valid for any a in a neighborhood of α so that for all a near α, $\widetilde{C} \cap C_a = \widetilde{C} \cap C_\alpha$. By Lemma 7, $\widetilde{C}$ is supported by a polytope P whose mixed volume with Q is d. □

5. Algorithm for toric absolute factorization

We describe an algorithm for the absolute factorization of a bivariate irreducible polynomial $f \in \mathbb{Q}[x]$ with Newton polytope P.

Let us denote by $X = \mathbb{T} \cup D_1 \cdots \cup D_r$ the abstract toric variety associated to P, where the divisor D_i corresponds to the facet P_i of P (see Appendix or Fulton (1993)). Assume that the origin is a vertex and that P_1 and P_2 contain it.

Algorithm. *Input:* A bivariate irreducible polynomial $f \in \mathbb{Q}[x]$.
Output: The absolute irreducible decomposition of f (i.e. its irreducible factorization in $\mathbb{C}[x]$).

(1) Determine the representation of P as intersection of affine half-planes:

$$P = \{m \in \mathbb{R}^2, \langle m, \eta_i\rangle + k_i \geq 0,\ i = 1 \dots r\}$$

such that $P_i = \{m \in P, \langle m, \eta_i\rangle + k_i = 0\}, i = 1 \dots r$, support the facets of P.

(2) Find the smallest integer polytope Q such that $P = dQ$, $d \in \mathbb{N}^*$. Let q be a generic Laurent polynomial supported by Q, and for $t \in \mathbb{C}$ generic, denote by $C_t \subset X$ the curve defined by $q(x) - t$. Determine the 0-cycle $C_t \cdot C = p_1(t) + \cdots + p_N(t)$ on X.

(3) For each $i = 3 \dots r$, determine the intersection set $C \cdot D_i = \{p_{i1}, \dots, p_{il_i}\}$ (each p_{ij} is repeated according to its multiplicity).

(4) Recognize a partition of $\{1, \dots, N\}$ (unique up to the labelling of multiple intersection points)

$$\mathcal{J} := (J_{31} \cup \cdots \cup J_{3l_3}) \cup \cdots \cup (J_{r1} \cup \cdots \cup J_{rl_r})$$

such that $\mathrm{card}(J_{ik}) = k_i' = \frac{k_i}{d} \in \mathbb{N}$ and $\lim_{|t|\to\infty} p_j(t) = p_{ik} \Longleftrightarrow j \in J_{ik}$.

(5) Find the biggest divisor δ of d such that for each $i = 3 \dots r$, there exists $J_i \subset \{1, \dots, l_i\}$ of cardinal $\frac{l_i}{\delta}$ satisfying

$$T_{\delta,J_3,\dots,J_r} := \sum_{i=3}^{r} \sum_{k\in J_i} \sum_{j\in J_{ik}} \frac{f_{x_2}}{\mathrm{Jac}(f, q)} \big(p_j(t)\big) = 0. \tag{9}$$

(6) Theorem 5 implies that f admits δ absolute irreducible factors whose traces on the facets $P_3, \dots, P_r$ are given by the partition $\mathcal{J}$. Make these factors explicit using Hensel's liftings as in Abu Salem et al. (2004) but with bigfloat coefficients as in Chèze and Galligo (2005) and Rupprecht (2004).

(7) From this approximate factorization, compute the extension $\mathbb{K}$ in Section 2 and recognize the exact factorization as explained in Chèze and Galligo (2005, 2006).

Remark 11. Let us comment some of these points.

Our main target is not polynomials with too small polytopes (which can be treated by other means), so we assume that $(1, 0)$ is not a vertex of Q.

The curve $C \subset X$ determined by f belongs to the linear system $|D_P| = |dD_Q|$, where $D_Q = k_3'D_3 + \cdots + k_r'D_r$.

The number of points N in the cycle $C_t \cdot C$ is equal, by Bernstein's theorem, to $d(D_Q \cdot D_Q) = 2d\mathrm{vol}(Q)$ (see Bernstein (1975)). The curve $C_t \subset X$ is the zero set of the homogeneous polynomial $Q^h(U) - t\prod_{i=3}^r U_i^{k_i'}$, where $U = (U_1, \dots, U_r)$ are homogeneous coordinates on X associated to the edges of Q and Q^h is the Q-homogenization of q (see Cox (1995)). When $|t|$ goes to infinity, C_t degenerates to the effective divisor at infinity $D_Q = \mathrm{div}_0(\prod_{i=3}^r U_i^{k_i'})$, and

$$p_1(t) + \cdots + p_N(t) \longrightarrow k_3'(p_{31} + \cdots + p_{3l_3}) + \cdots + k_r'(p_{r1} + \cdots + p_{rl_r}).$$

In the examples, to determine the partition of $\{1, \dots, N\}$ in the algorithm, we fix t with $|t|$ big and solve the polynomial system $f = q - t = 0$.

Proof of the algorithm. Let d' be a divisor of d and set $N' := \frac{N}{d'} = 2\frac{d}{d'}\mathrm{vol}(Q)$. To any subset $J = \{j_1, \dots, j_{N'}\}$ of $\{1, \dots, N\}$, we associate the 0-cycle

$$p_{j_1}(t) + \cdots + p_{j_{N'}}(t).$$

Since $(1,0) \notin \Gamma$ (Γ is defined in Theorem 9), and absolute irreducible factors of f are supported by a polytope homothetic to Q, the curve $C = \{f = 0\}$ intersects properly the Zariski closure of any line $x_1 = c$, $c \in \mathbb{C}$. Thus, Theorems 9 and 10 imply that there exists an algebraic curve $C_J \subset X$ such that for any $t \in \mathbb{C}$,

$$C_J \cdot C_t = p_{j_1}(t) + \cdots + p_{j_{N'}}(t)$$

if and only if the trace of x_1

$$T_J(t) := x_1(p_{j_1}(t)) + \cdots + x_1(p_{j_{N'}}(t))$$

does not depend on t. Such a curve is contained in C and is supported by $(d/d')Q$. If d' is the biggest divisor of d for which there exists a vanishing sum as in (9), $C_J = C_J(d')$ is an irreducible component of C, and f has d' irreducible factors.

Let us compute the finite sum $T_J = \sum_{j\in J} x_1\big(p_j(t)\big)$. The functions

$$u_j(t) = x_1(p_j(t)) \quad \text{and} \quad v_j(t) = x_2(p_j(t))$$

are holomorphic and satisfy for $j = 1 \dots N$,

$$f(u_j(t), v_j(t)) = 0, \qquad q(u_j(t), v_j(t)) = t.$$

Differentiating this system, we deduce that

$$u_j{}'(t) = -\frac{\partial_{x_2} f}{\mathrm{Jac}(f,q)}(p_j(t)), \qquad v_j{}'(t) = \frac{\partial_{x_1} f}{\mathrm{Jac}(f,q)}(p_j(t)).$$

Thus

$$T_J{}'(t) = -\sum_{j\in J} \frac{\partial_{x_2} f}{\mathrm{Jac}(f,q)}(p_j(t)).$$

The existence of the curve $C_J \subset C$ is then equivalent to $T_J{}'(t) = 0$ for q generic.

It remains to show the validity of step 4 in the algorithm. If C_J is a component of C, it has the same asymptotic behavior as that of C, i.e. the 0-cycle $C_t \cdot C_J$ converges to

$$D_Q \cdot C_J = k_3'(D_3 \cdot C_J) + \cdots + k_r'(D_r \cdot C_J).$$

The 0-cycle $C_t \cdot C_J = p_{j_1}(t) + \cdots + p_{j_{N'}}(t)$ is a sum of effective 0-cycles $Z_1(t), \dots, Z_r(t)$, where $Z_i(t)$ has degree $k_i' \frac{l_i}{d'}$ and $Z_i(t) \to k_i' D_i \cdot C_J$. □

5.1. Example

We apply our algorithm to the following simple (but not trivial) example:

$$\begin{aligned} f \;=\;& 49 + 30\,y\,x - 90\,y\,x^2 - 130\,x\,y^2 + 126\,y + 56\,x + 30\,x^2 - 3\,y^2 + x^4 + 8\,x^3 \\ &+ 36\,y^4 - 108\,y^3 - 127\,y^2\,x^2 + 32\,y^2\,x^3 - 54\,y\,x^3 + 84\,y^3\,x^2 + 37\,y^2\,x^4 \\ &- 12\,y\,x^4 + 30\,y^3\,x^3 + 13\,x^2\,y^4 + 24\,x\,y^4. \end{aligned}$$

The Newton polytope P of f represented in Fig. 3 is the convex hull of $\{(0,0), (4,0), (4,2), (2,4), (0,4)\}$. Also, $\eta_3 = (0,-1)$, $\eta_4 = (-1,-1)$, $\eta_5 = (-1,0)$, $k_3 = 4$, $k_4 = 6$, $k_5 = 4$, $d = 2$, and Q is the convex hull of $\{(0,0), (2,0), (2,1), (1,2), (0,2)\}$. Let

$$q = -5 + 8x - 2y + x^2 + y^2 + 2xy^2 + 6yx^2.$$

Fig. 4 illustrates the principle of our algorithm on this example.

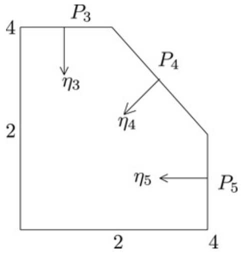

Fig. 3.

For $t = 10^3$, the intersection 0-cycle of the curve C_t defined by $q - t$ and the curve C defined by f is $C_t \cdot C = p_1 + \cdots + p_{14}$, with

$$
\begin{aligned}
p_1 &= (-3.788354357 - 22.18782564\,I,\ 0.1524031261 + 0.049759143\,I)\\
p_2 &= (-3.788354357 + 22.18782564\,I,\ 0.1524031261 - 0.049759143\,I)\\
p_3 &= (-2.389966107 - 4.663138871\,I,\ 7.365424369 + 1.227961352\,I)\\
p_4 &= (-2.389966107 + 4.663138871\,I,\ 7.365424369 - 1.227961352\,I)\\
p_5 &= (-1.986201832 - 22.37900395\,I,\ 0.1619217298 - 0.0018513709\,I)\\
p_6 &= (-1.986201832 + 22.37900395\,I,\ 0.1619217298 + 0.0018513709\,I)\\
p_7 &= (-1.535681765 - 1.726064601\,I,\ -9.102030424 + 7.399506679\,I)\\
p_8 &= (-1.535681765 + 1.726064601\,I,\ -9.102030424 - 7.399506679\,I)\\
p_9 &= (-1.045747272 - 3.489978116\,I,\ -5.189003901 + 9.662581013\,I)\\
p_{10} &= (-1.045747272 + 3.489978116\,I,\ -5.189003901 - 9.662581013\,I)\\
p_{11} &= (-0.7687604288 - 1.155834857\,I,\ 14.36735548 - 7.960507788\,I)\\
p_{12} &= (-0.7687604288 + 1.155834857\,I,\ 14.36735548 + 7.960507788\,I)\\
p_{13} &= (5.894022105 - 0.6210086653\,I,\ -6.648718394 + 5.938892046\,I)\\
p_{14} &= (5.894022105 + 0.6210086653\,I,\ -6.648718394 - 5.938892046\,I).
\end{aligned}
$$

Now to determine $C \cdot D_i$, $i = 3, 4, 5$, we use toric affine coordinates (see Appendix) to find the three facet polynomials of f. Using the chart corresponding to the vertex $s_3 = (2, 4)$ with the coordinates $u = \frac{1}{x}$, $v = \frac{x}{y}$, we find

$$f_3(u) = 36u^2 + 24u + 13 \quad \text{and} \quad f_4(v) = 37v^2 + 30v + 12.$$

In the chart associated to $s_4 = (2, 4)$ with the coordinates $z = \frac{y}{x}$, $w = \frac{1}{y}$, we obtain $f_5(w) = w^2 - 12w + 37$. So we have

$$C \cdot D_3 = \{p_{3,1}, p_{3,2}\}, \qquad C \cdot D_4 = \{p_{4,1}, p_{4,2}\}, \qquad C \cdot D_5 = \{p_{5,1}, p_{5,2}\},$$

where

$$
\begin{aligned}
&u(p_{3,1}) = -\frac{1}{3} + \frac{1}{2}I, \qquad v(p_{3,1}) = 0, \qquad u(p_{3,2}) = -\frac{1}{3} - \frac{1}{2}I, \qquad v(p_{3,2}) = 0,\\
&v(p_{4,1}) = -\frac{15}{37} + \frac{16}{37}I, \qquad u(p_{4,1}) = 0, \qquad v(p_{4,2}) = -\frac{15}{37} - \frac{16}{37}I, \qquad u(p_{4,2}) = 0,\\
&w(p_{5,1}) = 6 + I, \qquad z(p_{5,1}) = 0, \qquad w(p_{5,2}) = 6 - I, \qquad z(p_{5,2}) = 0,
\end{aligned}
$$

and

$$
\begin{aligned}
f_3(u) &= 36\big(u - u(p_{3,1})\big)\big(u - u(p_{3,2})\big),\\
f_4(v) &= 37\big(v - v(p_{4,1})\big)\big(v - v(p_{4,2})\big),\\
f_5(w) &= \big(w - w(p_{5,1})\big)\big(w - w(p_{5,2})\big).
\end{aligned}
$$

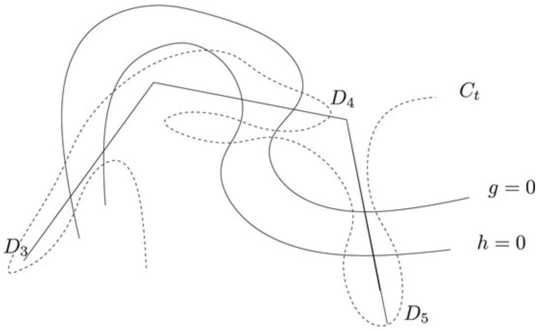

Fig. 4.

Now we collect the factors of f_i's to recover the factorization of f on the border $\Gamma = P_3 \cup P_4 \cup P_5$ of the Newton polytope P of f.

Since $C_t \cdot C = p_1(t) + \cdots + p_{14}(t)$, and $C_t \to 2D_3 + 3D_4 + 2D_5$, then 4 (resp. 6, and 4) points among these 14 converge to the 2 points in $C \cdot D_3$ (resp. $C \cdot D_4$, and $C \cdot D_5$), that is

$$p_1(t) + \cdots + p_{14}(t) \to 2(p_{3,1} + p_{3,2}) + 3(p_{4,1} + p_{4,2}) + 2(p_{5,1} + p_{5,2}).$$

More precisely, using the toric coordinates, we observe that the points p_1, p_6 (resp. p_3, p_{10}, p_{13}, and p_8, p_{12}) converge to $p_{5,1}$ (resp. $p_{4,1}$, and $p_{3,1}$). We deduce that

$$J_{3,1} = \{8, 12\}, \qquad J_{4,1} = \{3, 10, 13\}, \qquad J_{5,1} = \{1, 6\},$$
$$J_{3,2} = \{7, 11\}, \qquad J_{4,2} = \{4, 9, 14\}, \qquad J_{5,2} = \{13, 14\}.$$

Finally testing the vanishing of the expression (9), we find $\delta = 2, J_3 = \{1\}, J_4 = \{1\}, J_5 = \{1\}$. We deduce that the polynomial f admits 2 absolute irreducible factors g and h, and that the restriction of g on the 3 facets of P constituting Γ are (up to monomials)

$$g_3(u) = u - u(p_{3,1}), g_4(v) = v - v(p_{4,1}), g_5(w) = w - w(p_{5,1}).$$

We easily recognize the extension $\mathbb{K} = \mathbb{Q}[I]$ with $I^2 = 1$. In this extension, the polynomial coefficients are easily recognized from their decimal approximation.

Back to the toric coordinates (x, y), we find that the facet polynomials g_Γ (the restriction of g to Γ) and h_Γ are respectively

$$g_\Gamma = 6xy^2g_1 + xy^2g_3 = (2 - 3I)xy^2 + 6y^2 + (6 - I)x^2y - x^2,$$
$$h_\Gamma = (2 + 3I)xy^2 + 6y^2 + (6 + I)x^2y - x^2.$$

Remark 12. In this example to detect a partition of points defining the absolute factors of f we test $\binom{2}{1}\binom{2}{1}\binom{2}{1} = 6$ traces instead of $\binom{6}{3} = 20$ suggested by the original approach (see Section 3, Rupprecht (2004) and Chèze and Galligo (2005)). In general using our approach based on the partition given in the step 4 of the algorithm, we have to test at most

$$\mathcal{N} = \sum_{\delta | n} \prod_{i=1}^{r} \binom{\frac{e_i}{\delta}}{e_i}$$

traces instead of the initial number

$$\mathcal{M} = \sum_{\delta | n} \binom{\frac{d}{\delta}}{d}.$$

Since $d = e_1 + \cdots + e_r$ and

$$\binom{a}{b}\binom{c}{d} < \binom{a+c}{b+d},$$

this shows that $\mathcal{N} < \mathcal{M}$, and the difference being increasing with the number of facets of the Newton polytope of f. Our algorithmic approach will bring efficiency in absolute factorization problem and improves subsequently the approach presented in Rupprecht (2004) and Chèze and Galligo (2005).

5.2. Bidegree representation

In this subsection, the previous algorithmic approach and results are applied to the factorization of a polynomial f given by a dense representation of bidegree (m, n) in (x, y). Then the considered Newton polytope is simply a rectangle. Let C be the Zariski closure of the curve defined by f in the toric surface $\mathbb{P}^1 \times \mathbb{P}^1$. The algebricity criterion is expressed by cutting C with a family of conics having equations $xy + \alpha x + \beta y + \gamma = 0$ which can be written $(x - x_0)(y - y_0) - a = 0$. The trace of y is a rational function $\big(\mathrm{Tr}_C\big)$ of (x_0, y_0, a) defined in Definition 8 and the criterion is given in Theorem 9:

$$a \longmapsto \big(\mathrm{Tr}_C\, y\big)(a)$$

is polynomial of degree at most 1 in the coefficient a.

5.2.1. Translation

In order to apply this criterion, we perform a (generic) translation by (x_0, y_0) on the coordinates and get a new equation for f that we still denote by f to simplify the notation. So we intersect the curve C by a conic $xy = a$ with a near zero. Note a slight difference with the algorithm presented above, in this special case it is easier to cut the curve with lines and not with divisors at infinity, then a is considered near 0 and not near infinity. As (x_0, y_0) is generic when $a = 0$, there are m distinct intersection points denoted by $M_i = (x_i, 0)$ with the line $y = 0$ and n distinct intersection points denoted by $N_j = (0, y_i)$ with the line $x = 0$, moreover they are distinct from the origin. As a varies near zero, the intersection points form $m + n$ small curves which satisfy the equation in (a, y):

$$F(a, y) := y^m f(a/y, y) = 0.$$

Observe that $F(0, y) = 0$ has a root of multiplicity n. Hence our criterion can handle situations not covered by the criterion in Chèze and Galligo (2005).

5.2.2. Explicit criterion

Applications of the implicit function theorem for f at points M_i, respectively N_j, give the following Taylor expansions:

$$x = x_i + c_i y + O(y^2); \qquad y = y_j + a_j x + b_j x^2 + O(x^3)$$

where the numbers c_i, a_j, b_j are easily computed from the values of the first and second derivatives of f. Multiplying the first ones by y and the second ones by y, we get Taylor expansions with respect to a near M_i, respectively N_j :

$$y = \frac{1}{x_i}a - \frac{c_i}{x_i^3}a^2 + O(a^3); \qquad x = \frac{1}{y_j}a - \frac{a_j}{y_j^3}a^2 + O(a^3)$$

hence near N_j:

$$y = y_j + \frac{a_j}{y_j}a - \frac{a_j^2}{y_j^3}a^2 + \frac{b_j}{y_j^2}a^2 + O(a^3).$$

So the criterion becomes:

Explicit criterion: Two subsets of intersection points between C and the two axes, indexed by I of $\{1, \dots, m\}$ and J of $\{1, \dots, n\}$, correspond to an absolute irreducible component of C (hence to a factor of f) iff

$$\sum_{i \in I}\left(-\frac{c_i}{x_i^3}\right) + \sum_{j \in J}\left(-\frac{a_j^2}{y_j^3} + \frac{b_j}{y_j^2}\right) = 0.$$

5.2.3. Algorithm and comparison with total degree

Now, as in Chèze (2004), the LLL algorithm can be applied to solve the knapsack problem of size $m+n$ associated to the previous sums and determine the partitions of $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$ in q subsets that we denote by I_k and J_k with $k = 1, \ldots, q$. Each pair of such subsets should correspond to a factor f_k of f, a polynomial of bidegree $(\frac{m}{q}, \frac{n}{q})$. The solution of $f_k(x, 0) = 0$, respectively $f_k(0, y) = 0$, are the M_i indexed by I_k, respectively the N_j indexed by J_k. Moreover if we further assume that the constant term of f is 1 (which is easy to achieve), good approximations of $f_k(x, 0)$ and of $f_k(0, y)$ can be computed from the approximation of their roots.

Our algorithm consists in applying Hensel liftings with respect to x (respectively with respect to y) to lift the obtained approximate factorization of $f(0, y)$ (respectively of $f(x, 0)$) to an approximate factorization of f. We know that the factors must be conjugated; with the assumption we made on the constant term of f, their coefficients are conjugated algebraic integers. So an irreducible monic polynomial $g(z)$ defining a field extension can be recovered from a sufficiently good approximation by bigfloats of the coefficients as in Chèze and Galligo (2006). Then the exact expression of the coefficients can be recovered similarly.

Therefore the algorithm is completely similar to the one described in detail in Chèze and Galligo (2006) and the costs of the two approaches can be compared:

- Preprocessings (change of variables $f(x_0 + X, y_0 + Y)$ vs. $f(a_1Y + a_0, Y)$) have similar costs.
- Computations of traces and LLL also have similar costs.
- However, the new algorithm needs only m linear steps (or $\log(m)$ quadratic steps) of Hensel liftings of polynomials of degree n instead of $m+n$ linear steps (or $\log(m+n)$ quadratic steps) of Hensel liftings of polynomials of degree $m+n$.

  This makes a significant difference.

6. Conclusion

In this first paper, we established the mathematical bases of our algorithmic approach to toric factorization, and verified that it works on some examples. It is an important generalization of the algorithms developed by Rupprecht, Galligo and Chèze. We also presented in detail the case of a polynomial of bidegree (m, n) where we noted a significant improvement. However in the general case, we still have to tune and improve the algorithm. This will be done in a future work together with improvements which will speed it up in many cases of interest. The method is symbolic-numeric and produces approximate absolute factors to lift the approximate factorization to the exact one via rational approximation; we followed the model of computation used in Rupprecht (2004). Some additional work will also allow us to adapt the improvements of Chèze and Galligo (2006).

Let us for instance notice that we could replace the polytope $Q = \frac{1}{d}N_f$ by a smaller one $\tilde{Q}$ having parallel facets as we did in the bidegree case where we took $\tilde{Q}$ equal to the unit square.

We will also investigate the possibility of cutting the curve C defined by f by special families of curves which will ease the computations.

Acknowledgements

The authors would like to thank the anonymous referees for their comments and suggestions.

Appendix. Abstract toric surfaces

Let $Q \subset \mathbb{R}^2$ be a two-dimensional integer convex polytope satisfying the condition of Lemma 6. Let us explain how to recover the embedded projective toric variety X_Q as an abstract algebraic one.

There exist unique primitive vectors[1] $\eta_1, \ldots, \eta_r$ in $\mathbb{Z}^2$ and unique positive integers $k_1, \ldots, k_r$ in $\mathbb{N}$, such that for $i = 1 \ldots r$, the facet Q_i of Q is included in the affine line

$$Q_i \subset \{m \in \mathbb{R}^2, \langle m, \eta_i \rangle + k_i = 0\},$$

[1] A vector $v = (v_1, v_2) \in \mathbb{Z}^2$ is primitive if $\gcd(v_1, v_2) = 1$.

where $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^2$. The polytope Q is then given by the intersection of r affine half-planes:

$$Q = \{m \in \mathbb{R}^2 : \langle m, \eta_i \rangle + k_i \geq 0, \ \forall\, i = 1 \ldots r\}.$$

The vertices $s_1, \ldots, s_r$ of Q are in one-to-one correspondence with the facets of Q. If for $i = 1 \ldots r-1$, $s_i = Q_i \cap Q_{i+1}$, and $s_r = Q_r \cap Q_1$, any vertex s_i determines a two-dimensional rational convex cone

$$\sigma_i := \{m \in \mathbb{R}^2 : \langle m, \eta_i \rangle \geq 0, \langle m, \eta_{i+1} \rangle \geq 0\}$$

dual to the cone $\eta_i \mathbb{R}^+ \oplus \eta_{i+1} \mathbb{R}^+$. Let

$$X_i := \mathrm{Spec}(\mathbb{C}[\sigma_i \cap \mathbb{Z}^2])$$

be the biggest variety on which all the Laurent polynomials supported in σ_i can be extended as regular functions. Such a variety is called an affine toric surface, since the torus $\mathbb{T} = (\mathbb{C}^*)^2$ is an open set of X_i and its action on itself extends to X_i.

We can glue naturally the affine surfaces X_i and X_{i+1}, corresponding to cones having a common one-dimensional face, along their common set $X_i \cap X_{i+1}$ containing the torus $\mathbb{T}$. This natural gluing is compatible with the torus action and gives a complete normal variety X containing $\mathbb{T}$ as a Zariski open set. This torus compactification is called the normal complete toric surface associated to Q. It can be written as

$$X = \mathbb{T} \cup D_1 \cdots \cup D_r,$$

where $D_1, \ldots, D_r$ are the unique irreducible divisors of X invariant under the torus action. Each D_i is isomorphic to $\mathbb{P}^1$ and meets the affine toric variety X_k if and only if $k \in \{i, i+1\}$.

For any $m \in \mathbb{Z}^2$, the Laurent monomial x^m is regular on the Zariski open set $\mathbb{T}$ common to all the charts X_i. It defines a rational function on X giving rise to a principal Cartier divisor $\mathrm{div}(x^m)$ supported on $X \setminus \mathbb{T}$, and equal to

$$\mathrm{div}(x^m) = \sum_{i=1}^{r} \langle m, \eta_i \rangle \, D_i.$$

More generally, any Laurent polynomial q gives rise to a principal Cartier divisor

$$\mathrm{div}(f) = C_f - b_1 D_1 - \cdots - b_r D_r,$$

where C_f is the Zariski closure in X of the effective divisor $\{f = 0\} \subset \mathbb{T}$, and

$$b_i = -\min\{\langle m, \eta_i \rangle, m \in N_f\}, \quad i = 1 \ldots r,$$

are integers, N_f is the Newton polytope of f. Conversely, to any toric divisor $D = \sum_{i=1}^{r} b_i D_i$, we can associate an integral polytope P_D

$$P_D = \{m \in \mathbb{R}^2 : \langle m, \eta_i \rangle + b_i \geq 0, i = 1 \ldots r\}$$

so that $\mathrm{div}(f) + D \geq 0$ if and only if the support of f is contained in P_D, for any Laurent polynomial f. In other words, the set $H^0(X, \mathcal{O}_X(D))$ of global sections of the invertible sheaf corresponding to D is isomorphic to the set of Laurent polynomials supported by P_D, and admits the Laurent monomials x^m, $m \in P_D \cap \mathbb{Z}^2$, as a natural basis.

Let us denote by

$$D_Q = k_1 D_1 + \cdots + k_r D_r$$

the particular divisor associated to the given polytope Q (so that $Q = P_{D_Q}$). It is globally generated on X and gives rise to the Kodaira rational map

$$\phi_{D_Q} : X \longrightarrow \mathbb{P}(H^0(X, \mathcal{O}_X(D_Q)))^{\nu}$$

which sends a generic x on the point ζ_x corresponding to the hyperplane of global sections vanishing at x. If $x \in \mathbb{T}$, and $Q \cap \mathbb{Z}^2 = \{m_0, \dots, m_l\}$, this hyperplane is

$$\left\{a = [a_0 : \cdots : a_l] \in \mathbb{P}(H^0(X, \mathcal{O}_X(D))) : \sum_{i=0}^{l} a_i x^{m_i} = 0\right\}.$$

So that the natural homogeneous coordinates of ζ_x for $x \in \mathbb{T}$ are

$$\phi_{D_Q}(x) = \zeta_x = [x^{m_0} : \cdots : x^{m_l}],$$

and ϕ_{D_Q} defines a morphism on the torus. The map ϕ_{D_Q} turns out to be an embedding precisely when $m_1 - m_0, \dots, m_l - m_0$ generate the lattice $\mathbb{Z}^2$ (see Lemma 6), in this case the toric variety X is isomorphic to the projective variety X_Q previously constructed. The divisor D_Q is then very ample and gives rise to the isomorphism

$$H^0(X, \mathcal{O}_X(D_Q)) = \phi_{D_Q}^{*} H^0(\mathbb{P}^l, \mathcal{O}_{\mathbb{P}^l}(1)) \simeq H^0(X_Q, (\mathcal{O}_{\mathbb{P}^l}(1))_{|X_Q}), \tag{10}$$

traducing that the closure in X of curves defined by generic Laurent polynomials supported by Q are isomorphic to some hyperplane sections of $X_Q \subset \mathbb{P}^l$. We notice that the genericity criterion is essential here: For example, if $f(x) = x^{m_i}$, then the curve defined by f is empty while the corresponding hyperplane section $X_Q \cap \{u_i = 0\}$ is not. Let us explicit this genericity criterion.

Lemma 13. *Assume that D_Q is very ample and let f be a reduced Laurent polynomial supported in dQ, $d \in \mathbb{N}^*$. Then $C_f \simeq X_Q \cdot H$, for a reduced hypersurface $H \subset \mathbb{P}^l$ of degree d if and only if the support of f meets every facets of dQ.*

Proof. The assumption $N_f \subset dQ$ is equivalent to $\mathrm{div}(f) = C_f - D_f$, where $D_f = b_1 D_1 + \cdots + b_r D_r$ is an effective divisor bounded by dD_Q. Thus $\mathrm{div}(f) = C_f + (dD_Q - D_f) - dD_Q$, and since $C_f + (dD_Q - D_f) \geq 0$, f defines a global section of $\mathcal{O}_X(dD_Q)$. We deduce from the isomorphism (10), the existence of an effective divisor H of degree d in $\mathbb{P}^l$ such that

$$C_f + (dD_Q - D_f) = H_{|X},$$

under the identification $X = X_Q$. Then $C_f = H_{|X}$ if and only if $dD_Q = D_f$, that is if the equality $b_i = dk_i$ holds for every $i = 1 \dots r$. Moreover, as C_f is reduced, H must be reduced. □

References

Abu Salem, F., Gao, S., Lauder, A.G.B., 2004. Factoring polynomials via polytopes. In: ISSAC 2004. ACM, New York, pp. 4–11.

Bernstein, D. N., 1975. The number of roots of a system of equations. Funkcional. Anal. i Priložen. 9, 1–4.

Bostan, A., Lecerf, G., Salvy, B., Schost, E., Wiebelt, B., 2004. Complexity issues in bivariate polynomial factorization. In: ISSAC 2004. ACM, New York, pp. 42–49.

Bruns, W., Gubeladze, J., Ngô, V.T., 1997. Normal polytopes, triangulations, and Koszul algebras. J. Reine Angew. Math. 485, 123–160.

Chèze, G., 2004. Absolute polynomial factorization in two variables and the knapsack problem. In: ISSAC 2004. ACM, New York, pp. 87–94.

Chèze, G., Galligo, A., 2005. Four lectures on polynomial absolute factorization. In: Solving Polynomial Equations. In: Algorithms Comput. Math., vol. 14. Springer, Berlin, pp. 339–392.

Chèze, G., Galligo, A., 2006. From an approximate to an exact absolute polynomial factorization. J. Symbolic Comput. 41, 682–696.

Chèze, G., Lecerf, G., 2007. Lifting and recombination techniques for absolute factorization. J. Complexity 23, 380–420.

Cox, D.A., 1995. The homogeneous coordinate ring of a toric variety. J. Algebraic Geom. 4, 17–50.

Cox, D.A., 2003. What is a Toric Variety. In: Contemporary Mathematics, vol. 334. pp. 203–223.

Emiris, I.Z., 1996. On the complexity of sparse elimination. J. Complexity 12, 134–166.

Fulton, W., 1993. Introduction to toric varieties. In: The William H. Roever Lectures in Geometry. In: Annals of Mathematics Studies, vol. 131. Princeton University Press, Princeton, NJ.

Gao, S., 2001. Absolute irreducibility of polynomials via Newton polytopes. J. Algebra 237, 501–520.

Gao, S., 2003. Factoring multivariate polynomials via partial differential equations. Math. Comp. 72, 801–822 (electronic).

Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V., 1994. Discriminants, resultants, and multidimensional determinants. In: Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA.

Griffiths, P., Harris, J., 1978. Principles of algebraic geometry. In: Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York.

Hovanskiĭ, A.G., 1978. Newton polyhedra and the Euler–Jacobi formula. Uspekhi Mat. Nauk 33, 237–238.

Lecerf, G., 2007. Improved dense multivariate polynomial factorization algorithms. J. Symbolic Comput. 42, 477–494.

Newton, I., 1710. Curves. In: Lexicon Technicum, vol. 2.
Ostrowski, A.M., 1975. On multiplication and factorization of polynomials. I. Lexicographic orderings and extreme aggregates of terms. Aequationes Math. 13, 201–228.
Ruppert, W., 1986. Reduzibilität ebener Kurven. J. Reine Angew. Math. 369, 167–191.
Rupprecht, D., 2004. Semi-numerical absolute factorization of polynomials with integer coefficients. J. Symbolic Comput. 37, 557–574.
Sasaki, T., Suzuki, M., Kolář, M., Sasaki, M., 1991. Approximate factorization of multivariate polynomials and absolute irreducibility testing. Japan J. Indust. Appl. Math. 8, 357–375.
Sommese, A.J., Verschelde, J., Wampler, C.W., 2001. Numerical decomposition of the solution sets of polynomial systems into irreducible components. SIAM J. Numer. Anal. 38, 2022–2046 (electronic).
Sommese, A.J., Verschelde, J., Wampler, C.W., 2004. Numerical factorization of multivariate complex polynomials. Theoret. Comput. Sci. 315, 651–669.
von zur Gathen, J., Kaltofen, E., 1985. Factoring sparse multivariate polynomials. J. Comput. System Sci. 31, 265–287 (special issue: Twenty-fourth annual symposium on the foundations of computer science (Tucson, Ariz., 1983)).
Weimann, M., 2007. An interpolation theorem in toric varieties (submitted for publication).
Wood, J.A., 1984. A simple criterion for local hypersurfaces to be algebraic. Duke Math. J. 51, 235–237.