Change of basis for $\mathfrak{m}$-primary ideals in one and two variables

Seung Gyu Hyun
University of Waterloo
Waterloo, ON, Canada
kevin.hyun@edu.uwaterloo.ca

Stephen Melczer
University of Pennsylvania
Philadelphia, USA
smelczer@sas.upenn.edu

Éric Schost
University of Waterloo
Waterloo, ON, Canada
eschost@uwaterloo.ca

Catherine St-Pierre
University of Waterloo
Waterloo, ON, Canada
catherine.st-pierre@uwaterloo.ca

Abstract

Following recent work by van der Hoeven and Lecerf (ISSAC 2017), we discuss the complexity of linear mappings, called *untangling* and *tangling* by those authors, that arise in the context of computations with univariate polynomials. We give a slightly faster tangling algorithm and discuss new applications of these techniques. We show how to extend these ideas to bivariate settings, and use them to give bounds on the arithmetic complexity of certain algebras.

CCS Concepts

• **Computing methodologies → Symbolic and algebraic algorithms**;

Keywords

Polynomials; algorithms; complexity

ACM Reference Format:
Seung Gyu Hyun, Stephen Melczer, Éric Schost, and Catherine St-Pierre. 2019. Change of basis for $\mathfrak{m}$-primary ideals in one and two variables. In *2019 Int'l Symposium on Symbolic & Algebraic Computation (ISSAC'19), July 15–18, 2019, Beijing, China.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3326229.3326268

1 Introduction

In [22], van der Hoeven and Lecerf gave algorithms for "modular composition" modulo powers of polynomials: that is, computing $F(G) \bmod T^\mu$, for polynomials F, G, T over a field $\mathbb{F}$ and positive integer μ. As an intermediate result, they discuss a linear operation and its inverse, which they respectively call *untangling* and *tangling*.

Given separable $T \in \mathbb{F}[x]$ of degree d and a positive integer μ, polynomials modulo T^μ can naturally be written in the power basis $1, x, \dots, x^{d\mu-1}$. Here we consider another representation, based on bivariate polynomials. Introduce $\mathbb{K} := \mathbb{F}[y]/\langle T(y)\rangle$ with α the residue class of y; then, as an $\mathbb{F}$-algebra, $\mathbb{F}[x]/\langle T^\mu\rangle$ is isomorphic to $\mathbb{K}[\xi]/\langle \xi^\mu\rangle$ and untangling and tangling are the corresponding change of bases that maps x to $\xi + \alpha$. Take, for instance, $\mathbb{F} = \mathbb{Q}$, $T = x^2 + x + 2$ and $\mu = 2$. Then $\mathbb{K} = \mathbb{Q}[y]/\langle y^2 + y + 2\rangle$; untangling is the isomorphism $\mathbb{Q}[x]/\langle x^4 + 2x^3 + 5x^2 + 4x + 4\rangle \to \mathbb{K}[\xi]/\langle \xi^2\rangle$ and tangling is its inverse.

We now assume that $2, \dots, \mu - 1$ are units in $\mathbb{F}$. Van der Hoeven and Lecerf gave algorithms of quasi-linear cost for both untangling and tangling; their algorithm for tangling is slightly slower than that for untangling. Our first contribution is an improved algorithm for tangling, using duality techniques inspired by [33]. This saves logarithmic factors compared to the results in [22]; it may be minor in practice, but we believe this offers an interesting new point of view. Then we discuss how these techniques can be of further use, as in the resolution of systems of the form $F(x_1, x_2, x_3) = G(x_1, x_2, x_3) = 0$, for polynomials F, G in $\mathbb{F}[x_1, x_2, x_3]$.

Our second main contribution is an extension of these algorithms to situations involving more than one variable. As a first step, in this paper, we deal with certain systems in two variables. Indeed, the discussion in [22] is closely related to the question of how to describe isolated solutions of systems of polynomial equations. This latter question has been the subject of extensive work in the past; answers vary depending on what information one is interested in.

For the sake of this discussion, suppose we consider polynomials $G_1, \dots, G_s$ in the variables x_1 and x_2, with coefficients in $\mathbb{F}$. If one simply wants to describe set-theoretically the (finitely many) isolated solutions of $G_1, \dots, G_s$, popular choices include description by means of univariate polynomials [2, 9, 20, 27, 32], or triangular representations [3, 36]. When all isolated solutions are non-singular nothing else is needed, but further questions arise in the presence of multiple solutions as univariate or triangular representation may not be able to describe the local algebraic structure at such roots.

The presence of singular isolated solutions means that the ideal $\langle G_1, \dots, G_s\rangle$ admits a zero-dimensional primary component that is not radical. Thus, let I be a zero-dimensional primary ideal in $\mathbb{F}[x_1, x_2]$ with radical $\mathfrak{m}$; we will suppose that $\mathbb{F}[x_1, x_2]/\mathfrak{m}$ is separable (which is always the case if $\mathbb{F}$ is perfect, for instance) to prevent $\mathfrak{m}$ from acquiring multiple roots over $\overline{\mathbb{F}}$.

A direct approach to describing the solutions of I, together with the algebraic nature of I itself, is to give one of its Gröbner bases. Following [28], one may also give a basis of the dual of $\mathbb{F}[x_1, x_2]/I$, or a standard basis of I. In [28, Section 5], Marinari, Möller and Mora make the following interesting suggestion: build the field $\mathbb{K} := \mathbb{F}[y_1, y_2]/\tilde{\mathfrak{m}}$, where $\tilde{\mathfrak{m}}$ is the ideal $\mathfrak{m}$ with variables renamed y_1, y_2. Then the polynomials in I vanish at $\alpha := (\alpha_1, \alpha_2)$ when

ISSAC '19, July 15–18, 2019, Beijing, China

ACM ISBN 978-1-4503-6084-5/19/07…$15.00
https://doi.org/10.1145/3326229.3326268

α_1, α_2 are the residue classes of y_1, y_2 in $\mathbb{K}$. Now extend I to the polynomial ring $\mathbb{K}[\xi_1, \xi_2]$, for new variables ξ_1, ξ_2, by mapping (x_1, x_2) to (ξ_1, ξ_2). Then, the local structure of I at α can be described by the primary component of this extended ideal at α.

Let us show the similarities of this idea with van der Hoeven and Lecerf's approach, on an example from [30]. We take $\mathbb{F} = \mathbb{Q}$, $\mathfrak{m}$ is the maximal ideal $\langle T_1, T_2\rangle$, with $T_1 := x_1^2 + x_1 + 2$, $T_2 := x_2 - x_1 - 1$, and $I = \mathfrak{m}^2$ is the $\mathfrak{m}$-primary ideal with generators

$$G_3 = x_2^2 - 2x_1x_2 - 2x_2 + x_1^2 + 2x_1 + 1,$$
$$G_2 = x_1^2x_2 + x_1x_2 + 2x_2 - x_1^3 - 2x_1^2 - 3x_1 - 2,$$
$$G_1 = x_1^4 + 2x_1^3 + 5x_1^2 + 4x_1 + 4.$$

Since T_2 has degree one in x_2, we can simply take $\mathbb{K} := \mathbb{Q}[y_1]/\langle y_1^2 + y_1 + 2\rangle$, α_1 is the residue class of y_1 and $\alpha_2 = \alpha_1 + 1$.

The (α_1, α_2)-primary component J of the extension of I in $\mathbb{K}[\xi_1, \xi_2]$, i.e., the primary component associated to the prime ideal $(\xi_1 - \alpha_1, \xi_2 - \alpha_2)$, is the ideal with lexicographic Gröbner basis

$$H_3 = \xi_2^2 - 2\xi_2\alpha_1 - 2\xi_2 + \alpha_1 - 1,$$
$$H_2 = \xi_1\xi_2 - \xi_2\alpha_1 - \xi_1\alpha_1 - \xi_1 - 2,$$
$$H_1 = \xi_1^2 - 2\xi_1\alpha_1 - \alpha_1 - 2.$$

Its structure appears more clearly after applying the translation $(\xi_1, \xi_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$: the translated ideal J' admits the very simple Gröbner basis $\langle \xi_1^2, \xi_1\xi_2, \xi_2^2\rangle$. In other words, this representation allows one to complement the set-theoretic description of the solutions by the multiplicity structure.

Our first result in bivariate settings is the relation between the Gröbner bases of I and J (or J'): in our example, they both have three polynomials, and their leading terms are related by the transformation $(\xi_1, \xi_2) \mapsto (x_1^2, x_2)$. We then prove that, as in the univariate case, there is an $\mathbb{F}$-algebra isomorphism $\mathbb{F}[x_1, x_2]/I \to \mathbb{K}[\xi_1, \xi_2]/J'$ given by $(x_1, x_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$. In our example, this means that $\mathbb{Q}[x_1, x_2]/\langle G_1, G_2, G_3\rangle$ is isomorphic to $\mathbb{K}[\xi_1, \xi_2]/\langle \xi_1^2, \xi_1\xi_2, \xi_2^2\rangle$.

Under certain assumptions on J', we give algorithms for this isomorphism and its inverse that extend those for univariate polynomials; while their runtimes are not always quasi-linear, they are subquadratic in the degree of I (that is, the dimension of $\mathbb{F}[x_1, x_2]/I$). We end with a first application: upper bounds on the cost of arithmetic operations in an algebra such as $\mathbb{F}[x_1, x_2]/I$; these are new, to the best of our knowledge. Noteworthy that with a strong regularity assumption and a different setting, it has been shown in [35] that multiplication in $\mathbb{F}[x1, x2]/I$ can be done in quazi-linear time.

Although our results are still partial (we use assumptions and we only deal with bivariate systems), we believe it is worthwhile to investigate these questions. In future work, we plan to examine the impact of these techniques on issues arising from polynomial system solving algorithms: a direction that one may consider are lifting techniques in the presence of multiplicities, as in [21] for instance, as well as the computation of GCDs modulo ideals such as I above. See, for instance, [13] for a discussion of the latter question.

2 Preliminaries

In the rest of this paper, $\mathbb{F}$ is a *perfect* field. The costs of all our algorithms are measured in number of operations $(+, -, \times, \div)$ in $\mathbb{F}$.

2.1. We let $\mathsf{M} : \mathbb{N} \to \mathbb{N}$ be such that product of elements of degree less than n in $\mathbb{F}[x]$ can be computed in $\mathsf{M}(n)$ operations, and such that M satisfies the super-linearity properties of [17, Chapter 8]. Below, we will freely use all usual consequences of fast multiplication (on fast GCD, Newton iteration, ...) and refer the reader to e.g. [17] for details. In particular, multiplication in an $\mathbb{F}$-algebra of the form $\mathbb{A} := \mathbb{F}[x]/\langle T(x)\rangle$ with T monic in x, or $\mathbb{A} := \mathbb{F}[x_1, x_2]/\langle T_1(x_1), T_2(x_1, x_2)\rangle$ with T_1 monic in x_1 and T_2 monic in x_2, can be done in time $O(\mathsf{M}(\delta))$, with $\delta := \dim_{\mathbb{F}}(\mathbb{A})$. Inversion, when possible, is slower by a logarithmic factor. For $\mathbb{A} = \mathbb{F}[x_1, x_2]/I$, for a zero-dimensional monomial ideal I, multiplication and inversion in $\mathbb{A}$ can be done in time $O(\mathsf{M}(\delta)\log(\delta))$, resp. $O(\mathsf{M}(\delta)\log(\delta)^2)$, with $\delta = \dim_{\mathbb{F}}(\mathbb{A})$ (see the appendix).

2.2. We will use the *transposition principle* [10, 23], which is an algorithmic theorem stating that if the $\mathbb{F}$-linear map encoded by an $n \times m$ matrix over $\mathbb{F}$ can be computed in time T, the transposed map can be computed in time $T + O(n + m)$. This result has been used in a variety of contexts; our main sources of inspiration are [7, 33].

2.3. If $\mathbb{A}$ is an $\mathbb{F}$-vector space, its dual $\mathbb{A}^* := \mathrm{Hom}_{\mathbb{F}}(\mathbb{A}, \mathbb{F})$ is the $\mathbb{F}$-vector space of $\mathbb{F}$-linear mappings $\mathbb{A} \to \mathbb{F}$. When $\mathbb{A}$ is an $\mathbb{F}$-algebra, $\mathbb{A}^*$ becomes an $\mathbb{A}$-module: to a linear mapping $\ell : \mathbb{A} \to \mathbb{F}$ and $F \in \mathbb{A}$ we can associate the linear mapping $F \cdot \ell : G \in \mathbb{A} \mapsto \ell(FG)$. This operation is called the *transposed product* in $\mathbb{A}^*$, since it is the transpose of the multiplication-by-F mapping.

Given a basis $\mathcal{B}$ of $\mathbb{A}$, elements of $\mathbb{A}^*$ are represented on the dual basis, by their values on $\mathcal{B}$. In terms of complexity, if $\mathbb{A}$ is an algebra such as those in **2.1**, the transposition principle implies that transposed products can be done in time $O(\mathsf{M}(\delta))$, resp. $O(\mathsf{M}(\delta)\log(\delta))$, with again $\delta := \dim_{\mathbb{F}}(\mathbb{A})$. See [34] for detailed algorithms in the cases $\mathbb{A} = \mathbb{F}[x]/\langle T(x)\rangle$ and $\mathbb{A} = \mathbb{F}[x_1, x_2]/\langle T_1(x_1), T_2(x_1, x_2)\rangle$.

An element $\ell \in \mathbb{A}^*$ is called a *generator* of $\mathbb{A}^*$ if $\mathbb{A} \cdot \ell = \mathbb{A}^*$ (in other words, for any ℓ' in $\mathbb{A}^*$ there exists $F \in \mathbb{A}$, which must be unique, such that $F \cdot \ell = \ell'$). When $\mathbb{A} = \mathbb{F}[x]/\langle T(x)\rangle$, with $n := \deg(T)$, ℓ defined by $\ell(1) = \cdots = \ell(x^{n-2}) = 0$ and $\ell(x^{n-1}) = 1$ is known to generate $\mathbb{A}^*$. For $\mathbb{A} = \mathbb{F}[x_1, x_2]/\langle T_1(x_1), T_2(x_1, x_2)\rangle$, ℓ given by $\ell(x_1^{n_1-1}x_2^{n_2-1}) = 1$, with all other $\ell(x_1^i x_2^j) = 0$, is a generator (here, we write $n_1 := \deg(T_1, x_1)$ and $n_2 := \deg(T_2, x_2)$). For more general $\mathbb{A}$, $\mathbb{A}^*$ may not be free: see for example Subsection 4.4.

3 The univariate case revisited

In this section, we work with univariate polynomials. Considering that $T \in \mathbb{F}[x]$ is a monique and separable (that is, without roots repeated in $\overline{\mathbb{F}}$) with degree d, and letting μ be an integer positive, we start from the following hypothesis:

$\mathbf{H}_1$. $\mathbb{F}$ has characteristic at least μ.

Define $\mathbb{K} := \mathbb{F}[y]/T(y)$, and let α be the residue class of y in $\mathbb{K}$. Van der Hoeven and Lecerf proved that the $\mathbb{F}$-algebra mapping

$$\begin{array}{rccc} \pi_{T,\mu}: & \mathbb{F}[x]/\langle T^\mu\rangle & \to & \mathbb{K}[\xi]/\langle \xi^\mu\rangle \\ & x & \mapsto & \xi + \alpha \end{array}$$

is well-defined and realizes an isomorphism of $\mathbb{F}$-algebras. The mapping $\pi_{T,\mu}$ is called *untangling*, and its inverse $\pi_{T,\mu}{}^{-1}$ *tangling*. Note that $\pi_{T,\mu}(F)$ simply computes the first μ terms of the Taylor expansion of F at α, that is, $\pi_{T,\mu}(F) = \sum_{0 \le i < \mu} F^{(i)}(\alpha)\xi^i/i!$.

Reference [22] gives algorithms for both untangling and tangling, the latter calling the former recursively; their untangling algorithm runs in $O(\mathsf{M}(d\mu)\log(\mu))$ operations in $\mathbb{F}$, while their tangling algorithm takes $O(\mathsf{M}(d\mu)\log(\mu)^2 + \mathsf{M}(d)\log(d))$ operations. Using transposition techniques from [33], we prove the following.

Proposition 3.1. *Given G in $\mathbb{K}[\xi]/\langle\xi^\mu\rangle$, one can compute ${\pi_{T,\mu}}^{-1}(G)$ in $O(\mathsf{M}(d\mu)\log(\mu) + \mathsf{M}(d)\log(d))$ operations in $\mathbb{F}$.*

The $\mathbb{F}$-algebra $\mathbb{K}$ admits the basis $(1,\ldots,\alpha^{d-1})$; $\mathbb{F}[x]/\langle T^\mu\rangle$ has basis $\mathcal{B} = (1, x, \ldots, x^{d\mu-1})$ and $\mathbb{K}[\xi]/\langle\xi^\mu\rangle$ admits the bivariate basis $C = (1,\ldots,\alpha^{d-1},\xi,\ldots,\alpha^{d-1}\xi,\ldots\xi^{\mu-1},\ldots,\alpha^{d-1}\xi^{\mu-1})$. As per **2.3**, we represent a linear form $L \in \mathbb{F}[x]/\langle T^\mu\rangle^*$ by the vector $[L(x^i) \mid 0 \le i < d\mu] \in \mathbb{F}^{d\mu}$, and a linear form $\ell \in \mathbb{K}[\xi]/\langle\xi^\mu\rangle^*$ by the bidimensional vector $[\ell(\alpha^i\xi^j) \mid 0 \le i < d, 0 \le j < \mu] \in \mathbb{F}^{d\times\mu}$.

3.1 A faster tangling algorithm

This section shows that using the transpose of untangling allows us to deduce an algorithm for tangling; see [14, 33] for a similar use of transposition techniques. We start by describing useful subroutines.

3.1.1. The first algorithmic result we will need concerns the cost of inversion in $\mathbb{F}[x]/\langle T^\mu\rangle$. To compute $1/F \bmod T^\mu$ for some $F \in \mathbb{F}[x]$ of degree less than $d\mu$ we may start by computing $\bar{G} := 1/\bar{F} \bmod T$, with $\bar{F} := F \bmod T$; this costs $O(\mathsf{M}(d\mu) + \mathsf{M}(d)\log(d))$ operations in $\mathbb{F}$. Then we lift $\bar{G}$ to $G := 1/F \bmod T^\mu$ by Newton iteration modulo the powers of T, at the cost of another $O(\mathsf{M}(d\mu))$.

3.1.2. Next, we discuss the solution of certain Hankel systems. Consider L and L' two $\mathbb{F}$-linear forms $\mathbb{F}[x]/\langle T^\mu\rangle \to \mathbb{F}$; our goal is to find F in $\mathbb{F}[x]/\langle T^\mu\rangle$ such that $F \cdot L = L'$, under the assumption that L generates the dual space $\mathbb{F}[x]/\langle T^\mu\rangle^*$. In matrix terms, this is equivalent to finding coefficients $f_0,\ldots,f_{d\mu-1}$ of F such that $[H][f_0,\ldots,f_{d\mu-1}]^T = [f_0,\ldots,f_{d\mu-1}]^T$ with $H_{i,j} = L(x^{i+j})$. The system can be solved in $O(\mathsf{M}(d\mu)\log(d\mu))$ operations in $\mathbb{F}$ [8], but we may derive an improvement from the fact that T^μ is a μth power.

An algorithm that realizes the transposed product $(L,F) \mapsto L'$ is in [6, Lemma 2.5]: let $\zeta : \mathbb{F}^{d\mu} \to \mathbb{F}^{d\mu}$ be the upper triangular Hankel operator with first column the coefficients of degree $1,\ldots,d\mu$ of T^μ, and let Λ and Λ' be the two polynomials in $\mathbb{F}[x]$ with respective coefficients $\zeta(L)$ and $\zeta(L')$. Then $\Lambda' = F\Lambda \bmod T^\mu$.

Given the values of L and L' at $1,\ldots,x^{d\mu-1}$, we compute $\zeta(L)$ and $\zeta(L')$ in $O(\mathsf{M}(d\mu))$ operations. Since L generates $\mathbb{F}[x]/\langle T^\mu\rangle^*$, Λ is invertible modulo T^μ; then, using **3.1.1**, we compute its inverse in $O(\mathsf{M}(d\mu) + \mathsf{M}(d)\log(d))$ operations. Multiplication by Λ' takes another $O(\mathsf{M}(d\mu))$ operations, for a total of $O(\mathsf{M}(d\mu) + \mathsf{M}(d)\log(d))$.

3.1.3. We now recall van der Hoeven and Lecerf's algorithm for the mapping $\pi_{T,\mu}$, and deduce an algorithm for its transpose, with the same asymptotic runtime. Van der Hoeven and Lecerf's algorithm is recursive, with a divide-and-conquer structure; the key idea is that the coefficients of $\pi_{T,\mu}(F)$, for F in $\mathbb{F}[x]/\langle T^\mu\rangle$, are the values of $F, F', \ldots, F^{(\mu-1)}$ at α, divided respectively by $0!, 1!, \ldots, (\mu-1)!$.

Algorithm 1 $\pi_{\mathrm{rec}}(F, T, \mu)$

Input: $F \in \mathbb{F}[x]/\langle T^\mu\rangle$
Output: $[F(\alpha),\ldots,F^{(\mu-1)}(\alpha)] \in \mathbb{K}^\mu$
if $\mu = 1$ **then return** $[\,F(\alpha)\,]$ **else** set $\lambda := \lfloor \frac{\mu}{2} \rfloor$
return $\pi_{\mathrm{rec}}(F \bmod T^\lambda, T, \lambda)$ cat $\pi_{\mathrm{rec}}(F^{(\lambda)} \bmod T^{\mu-\lambda}, T, \mu-\lambda)$

Algorithm 2 $\pi(F, T, \mu)$

Input: $F \in \mathbb{F}[x]/\langle T^\mu\rangle$
Output: $\pi_{T,\mu}(F) \in \mathbb{K}[\xi]/\langle\xi^\mu\rangle$
return $\sum_{0\le i<\mu} \frac{v[i]}{i!}\xi^i$, with $v := \pi_{\mathrm{rec}}(F, T, \mu)$

The runtime $T(d,\mu)$ of π_{rec} satisfies $T(d,\mu) \le T(d,\mu/2) + O(\mathsf{M}(d\mu))$, so this results in an algorithm for $\pi_{T,\mu}$ that takes $O(\mathsf{M}(d\mu)\log(\mu))$ operations. Since $\pi_{T,\mu}$ is an $\mathbb{F}$-linear mapping $\mathbb{F}[x]/\langle T^\mu\rangle \to \mathbb{K}[\xi]/\langle\xi^\mu\rangle$, its transpose ${\pi_{T,\mu}}^\perp$ is an $\mathbb{F}$-linear mapping $\mathbb{K}[\xi]/\langle\xi^\mu\rangle^* \to \mathbb{F}[x]/\langle T^\mu\rangle^*$. The transposition principle implies that ${\pi_{T,\mu}}^\perp$ can be computed in $O(\mathsf{M}(d\mu)\log(\mu))$ operations; we make the corresponding algorithm explicit as follows.

We transpose all steps of the algorithm above, in reverse order. As input we take $\ell \in \mathbb{K}[\xi]/\langle\xi^\mu\rangle^*$, which we see as a bidimensional vector in $\mathbb{F}^{d\times\mu}$; we also write $\ell = [\ell_i \mid 0 \le i < \mu]$, with all ℓ_i in $\mathbb{F}^d$. The transpose of the concatenation at the last step allows one to apply the two recursive calls to the first and second halves of input ℓ. Each of them is followed by an application of the transpose of Euclidean division (see below), and after "transpose differentiating" the second intermediate result (see below), we return their sum.

Algorithm 3 $\pi_{\mathrm{rec}}^\perp(\ell, T, \mu)$

Input: $\ell \in \mathbb{F}^{d\times\mu}$
if $\mu = 1$ **then return** ℓ_0 **else** $\lambda := \lfloor \frac{\mu}{2} \rfloor$
$v_0 := \pi_{\mathrm{rec}}^\perp([\ell_i \mid 0 \le i < \lambda], T, \lambda)$ and $u_0 := \mathrm{mod}^\perp(v_0, T^\lambda, d\mu)$
$v_1 := \pi_{\mathrm{rec}}^\perp([\ell_i \mid \lambda \le i < \mu], T, \mu-\lambda)$
$u_1 := \mathrm{diff}^\perp(\mathrm{mod}^\perp(v_1, T^{\mu-\lambda}, d\mu-\lambda)), \lambda)$
return $u_0 + u_1$

Algorithm 4 $\pi^\perp(\ell, T, \mu)$

Input: $\ell \in \mathbb{K}[\xi]/\langle\xi^\mu\rangle^* \simeq \mathbb{F}^{d\times\mu}$
Output: ${\pi_{T,\mu}}^\perp(\ell) \in \mathbb{F}[x]/\langle T^\mu\rangle^* \simeq \mathbb{F}^{d\mu}$
return $\pi_{\mathrm{rec}}^\perp([\ell_i/i! \mid 0 \le i < \mu], T, \mu)$

Correctness follows from the correctness of van der Hoeven and Lecerf's algorithm. Following [7], given a vector u, a polynomial $S \in \mathbb{F}[x]$ and an integer $t \ge \deg(S)$, where u has length $\deg(S)$, $\mathrm{mod}^\perp(u, S, t)$ returns the first t terms of the sequence defined by initial conditions u and minimal polynomial S in time $O(\mathsf{M}(t))$. Given a vector u of length $t-\lambda$, $v := \mathrm{diff}^\perp(u,\lambda)$ is the vector of length t given by $v_0 = \cdots = v_{\lambda-1} = 0$ and $v_i = i\cdots(i-\lambda+1)u_{i-\lambda}$ for $i = \lambda,\ldots,t-1$. It can be computed in linear time $O(t)$. Overall, as in [22], the runtime is $O(\mathsf{M}(d\mu)\log(\mu))$.

3.1.4. We can now give our algorithm for the tangling operator ${\pi_{T,\mu}}^{-1}$; it is inspired by a similar result due to Shoup [33].

Take G in $\mathbb{K}[\xi]/\langle\xi^\mu\rangle$: we want to find $F \in \mathbb{F}[x]/\langle T^\mu\rangle$ such that $\pi_{T,\mu}(F) = G$. Let $\ell : \mathbb{K}[\xi]/\langle\xi^\mu\rangle \to \mathbb{F}$ be defined by $\ell(\alpha^{d-1}\xi^{\mu-1}) = 1$ and $\ell(\alpha^i\xi^j) = 0$ for all other values of $i < d, j < \mu$; as pointed out in **2.3**, this is a generator of $\mathbb{K}[\xi]/\langle\xi^\mu\rangle^*$. Define further $\ell' := G \cdot \ell$. Then ℓ' is a transposed product as in **2.3**, and we saw that it can be computed in $O(\mathsf{M}(d\mu))$ operations. This implies $\pi_{T,\mu}(F) \cdot \ell = \ell'$.

Let now $L := {\pi_{T,\mu}}^\perp(\ell)$ and $L' := {\pi_{T,\mu}}^\perp(\ell')$; we obtain them by applying our transpose untangling algorithm to ℓ, resp. ℓ', in time $O(\mathsf{M}(d\mu)\log(\mu) + \mathsf{M}(d)\log(d))$. Since ℓ is a generator of $\mathbb{K}[\xi]/\langle\xi^\mu\rangle^*$, L is a generator of $\mathbb{F}[x]/\langle T^\mu\rangle^*$. The equation $\pi_{T,\mu}(F) \cdot \ell = \ell'$ then implies that $F \cdot L = L'$, which is an instance of the problem discussed in **3.1.2**; applying the algorithm there takes another $O(\mathsf{M}(d\mu) + \mathsf{M}(d)\log(d))$. Summing all costs, this gives an algorithm for ${\pi_{T,\mu}}^{-1}$ with cost $O(\mathsf{M}(d\mu)\log(\mu) + \mathsf{M}(d)\log(d))$, proving Proposition 3.1.

3.2 Applications

3.2.1. For P in $\mathbb{F}[x]$ one can compute x^D mod P using $O(\log(D))$ multiplications modulo P by repeated squaring. Applications include Fiduccia's algorithm for the computation of terms in linearly recurrent sequences [16] or of high powers of matrices [19, 31]. This algorithm takes $O(\mathsf{M}(n)\log(D))$ operations in $\mathbb{F}$, with $n := \deg(P)$. We assume without loss of generality that $D \ge n$.

We can do better, in cases where P is not squarefree. For computations of terms in recurrent sequences, such P's appear when computing terms of *bivariate* recurrent sequences $(a_{i,j})$ defined by $\sum_{i,j} a_{i,j}x^iy^j = N(x,y)/Q(x,y)$, for some polynomials $N, Q \in \mathbb{F}[x,y]$ with $Q(0,0) \neq 0$. Then, the j-th row $\sum_i a_{i,j}x^i$ has characteristic polynomial P^j, where P is the reverse polynomial of $Q(x,0)$ [5].

First, assume that $P = T^\mu$ with T separable of degree d. Then we compute x^D mod P by tangling $r := (\xi + \alpha)^D$. The quantity $r = \sum_{i=0}^{\mu-1} \binom{D}{i} \xi^i \alpha^{D-i}$ can be computed in time $O(\mathsf{M}(d)(\log(D) + \mu))$, by computing $\alpha^{D-\mu+1}, \alpha^{D-\mu+2}, \ldots, \alpha^D$ and multiplying them by the binomial coefficients (which themselves are obtained by using the recurrence they satisfy). By Proposition 3.1, the cost of tangling is $O(\mathsf{M}(d\mu)\log(\mu) + \mathsf{M}(d)\log(d))$, which brings the total to $O(\mathsf{M}(d)\log(D) + \mathsf{M}(d\mu)\log(\mu))$, since $d \le D$. To compute x^D modulo an arbitrary P, one may compute the squarefree decomposition of P, apply the previous algorithm modulo each factor and obtain the result by applying the Chinese Remainder Theorem. The overall runtime becomes $O(\mathsf{M}(m)\log(D) + \mathsf{M}(n)\log(n))$, where n and m are the degrees of P and its squarefree part, respectively; this is to be compared with the cost $O(\mathsf{M}(n)\log(D))$ of repeated squaring. While this algorithm improves over the direct approach, practical gains show up for astronomical values of the parameters.

3.2.2. Assume $\mathbb{F} = \mathbb{Q}$. In [26], Lebreton, Mehrabi and Schost gave an algorithm to compute the intersection of surfaces in 3d-space, that is, to solve polynomial systems of the form $F(x_1,x_2,x_3) = G(x_1,x_2,x_3) = 0$. Assuming that the ideal $K := \langle F, G\rangle \subset \mathbb{Q}(x_1)[x_2,x_3]$ is radical and that we are in generic coordinates, the output is polynomials S, T, U in $\mathbb{Q}[x_1,x_2]$ such that K is equal to $\langle S, Ux_3 - T\rangle$ (so S describes the projection of the common zeros of F and G on the x_1,x_2-plane, and T and U allow us to recover x_3). The algorithm of [26] is Monte Carlo, with runtime $O(D^{4.7})$ where D is an upper bound on $\deg(F)$ and $\deg(G)$. The output has $\Theta(D^4)$ terms in the worst case, and the result in [26] is the best to date.

The case of non-radical systems was discussed in [29]. It was pointed out in the introduction of that paper that quasi-linear time algorithms for untangling and tangling (which were not explicitly called by these names) would make it possible to extend the results of [26] to general systems. Hence, already with the results by van der Hoeven and Lecerf a runtime $O(D^{4.7})$ was made possible for the problem of surface intersection, without radicality assumption.

4 The bivariate case

We now generalize the previous questions to the bivariate settings. We expect several of these ideas to carry over to higher numbers of variables, but some adaptations may be non-trivial (for instance, we rely on Lazard's structure theorem on lexicographic bivariate Gröbner bases). As an application, we give results on the complexity of arithmetic modulo certain primary ideals.

4.1 Setup

4.1.1. For the rest of the paper, the *degree* $\deg(I)$ of a zero-dimensional ideal I in $\mathbb{F}[x_1,x_2]$ is defined as the dimension of $\mathbb{F}[x_1,x_2]/I$ as a vector space (the same definition will hold for polynomials over any field).

Let $\mathfrak{m}$ be a maximal ideal of degree d in $\mathbb{F}[x_1,x_2]$; we consider two new variables y_1, y_2, we let $\gamma : \mathbb{F}[x_1,x_2] \to \mathbb{F}[y_1,y_2]$ be the $\mathbb{K}$-algebra isomorphism mapping (x_1,x_2) to (y_1,y_2) and let $\tilde{\mathfrak{m}} := \gamma(\mathfrak{m})$. This is a maximal ideal as well, and $\mathbb{K} := \mathbb{F}[y_1,y_2]/\tilde{\mathfrak{m}}$ is a field extension of degree d of $\mathbb{F}$. We then let α_1, α_2 be the respective residue classes of y_1, y_2 in $\mathbb{K}$.

Next, let $J \subset \mathbb{K}[\xi_1,\xi_2]$, for two new variables ξ_1, ξ_2, be a zero-dimensional primary ideal at $\alpha := (\alpha_1,\alpha_2)$. Finally, let $I := \Phi^{-1}(J)$, where Φ is the natural embedding $\mathbb{F}[x_1,x_2] \to \mathbb{K}[\xi_1,\xi_2]$ given by $(x_1,x_2) \mapsto (\xi_1,\xi_2)$. One easily checks that I is $\mathfrak{m}$-primary (that is, $\mathfrak{m}$ is the radical of I), and that J is the primary component at α of the ideal $I \cdot \mathbb{K}[\xi_1,\xi_2]$ generated by $\Phi(I)$. Note that since $\mathbb{F}$ is perfect, $\mathbb{F} \to \mathbb{K}$ is separable, so over an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$, $\mathfrak{m}$ has d distinct solutions. We make the following assumption:

H₂. $\mathbb{F}$ has characteristic at least n, with $n := \deg(I)$.

Finally, we let $J' \subset \mathbb{K}[\xi_1,\xi_2]$ be the ideal obtained by applying the translation $(\xi_1,\xi_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$ to J; it is primary at $(0,0)$.

4.1.2. Although our construction starts from the datum of $\mathfrak{m}$ and $J \subset \mathbb{K}[\xi_1,\xi_2]$ and defines I from them, we may also take as starting points $\mathfrak{m}$ and an $\mathfrak{m}$-primary ideal $I \subset \mathbb{F}[x_1,x_2]$ (this is what we did for the example in the introduction).

Under that point of view, consider the ideal $I \cdot \mathbb{K}[\xi_1,\xi_2]$ generated by $\Phi(I)$, for $\Phi : \mathbb{F}[x_1,x_2] \to \mathbb{K}[\xi_1,\xi_2]$ as above, and let J be the primary component of $I \cdot \mathbb{K}[\xi_1,\xi_2]$ at α. One verifies that I is equal to $\Phi^{-1}(J)$, so we are indeed in the same situation as in **4.1.1**.

4.1.3. For the rest of the paper, we use the lexicographic monomial ordering in $\mathbb{F}[x_1,x_2]$ induced by $x_1 < x_2$, and its analogue in $\mathbb{K}[\xi_1,\xi_2]$; "the" Gröbner basis of an ideal is its minimal reduced Gröbner basis for this order. Our first goal in this section is then to describe the relation between the Gröbner bases of I and J: viz., they have the same number of polynomials, and their leading terms are related in a simple fashion (as seen on the example).

Let T be the Gröbner basis of $\mathfrak{m}$. Since $\mathfrak{m}$ is maximal, T consists of two polynomials (T_1,T_2), with T_1 of degree d_1 in $\mathbb{F}[x_1]$ and T_2 in $\mathbb{F}[x_1,x_2]$, monic of degree d_2 in x_2. Note that $d_1d_2 = d = \deg(\mathfrak{m})$. Next, let $H = (H_1,\ldots,H_t)$ be the Gröbner basis of J, with $H_1 < \cdots < H_t$; we let $\xi_1^{\mu_1}\xi_2^{\nu_1}, \ldots, \xi_1^{\mu_t}\xi_2^{\nu_t}$ be the respective leading terms of $H_1,\ldots,H_t$. Thus, the μ_i's are decreasing, the ν_i's are increasing, and $\nu_1 = \mu_t = 0$. Finally, we let $\mu := \deg(J) = \deg(J')$. Remark that the Gröbner basis of J' admits the same leading terms as H.

In our example, we have $t = 3$, $(\mu_1,\nu_1) = (2,0)$, $(\mu_2,\nu_2) = (1,1)$ and $(\mu_3,\nu_3) = (0,2)$. The integers d_1, d_2 are respectively 2 and 1, so $d = 2$, the degree n is 6 and the multiplicity μ is 3. The key result in this subsection is the following.

Proposition 4.1. *The Gröbner basis of I has the form $(R_1,\ldots,R_t)$, where for $j = 1,\ldots,t$, $R_j = T_1^{\mu_j}\tilde{R}_j$, for some polynomial $\tilde{R}_j \in \mathbb{F}[x_1,x_2]$ monic of degree $d_2\nu_j$ in x_2. In particular, $n = d\mu$.*

As a result, for all j the leading term of R_j is ${x_1}^{d_1\mu_j}{x_2}^{d_2\nu_j}$, whereas that of H_j is ${\xi_1}^{\mu_j}{\xi_2}^{\nu_j}$, as in our example. The next two paragraphs are devoted to the proof of this proposition.

4.1.4. We define here a family of polynomials $G_1, \dots, G_t$, and prove that they form a (non-reduced) Gröbner basis of I in **4.1.5**.

Because the extension $\mathbb{F} \to \mathbb{K}$ is separable, it admits a primitive element β, with minimal polynomial $F \in \mathbb{F}[t]$; this polynomial has degree $[\mathbb{K} : \mathbb{F}] = d$. Let $\mathbb{L}$ be a splitting field for F containing $\mathbb{K}$ and let $I \cdot \mathbb{L}[\xi_1, \xi_2]$ and K be the extensions of $I \cdot \mathbb{K}[\xi_1, \xi_2]$ and J in $\mathbb{L}[\xi_1, \xi_2]$, respectively. Then $\deg(J) = \deg(K)$, and K is the primary component of $I \cdot \mathbb{L}[\xi_1, \xi_2]$ at α.

Let $\beta_1 = \beta, \beta_2, \dots, \beta_d$ be the roots of F in $\mathbb{L}$. For all $i = 1, \dots, d$, we let σ_i be an element in the Galois group of $\mathbb{L}/\mathbb{F}$ such that $\beta_i = \sigma_i(\beta)$, as well as $\alpha^{(i)} := (\sigma_i(\alpha_1), \sigma_i(\alpha_2))$. Note that these elements are pairwise distinct: since β is in $\mathbb{F}[\alpha_1, \alpha_2]$ and all σ_i's fix $\mathbb{F}$, $\alpha^{(i)} = \alpha^{(j)}$ implies $\beta_i = \beta_j$, and thus $i = j$. Therefore, $\alpha^{(1)}, \dots, \alpha^{(d)}$ can be seen as all the roots of $\mathfrak{m}$, with $\alpha^{(1)} = \alpha$.

For $i = 1, \dots, d$, let K_i be the primary component of $I \cdot \mathbb{L}[\xi_1, \xi_2]$ at $\alpha^{(i)}$, so that $K_1 = K$. By construction, these ideals are pairwise coprime, and their product is $I \cdot \mathbb{L}[\xi_1, \xi_2]$. Take i in $1, \dots, d$, and let D be a large enough integer such that $K = I \cdot \mathbb{L}[\xi_1, \xi_2] + \mathfrak{n}^D$ and $K_i = I \cdot \mathbb{L}[\xi_1, \xi_2] + \mathfrak{n}_i^D$, with $\mathfrak{n}$ and $\mathfrak{n}_i$ the maximal ideals at α and $\alpha^{(i)}$ respectively. Since $I \cdot \mathbb{L}[\xi_1, \xi_2]$ is defined over $\mathbb{F}$, σ_i thus maps the generators of K to those of K_i. This implies that the Gröbner basis of K_i is $(H_{i,1}, \dots, H_{i,t})$, with $H_{i,j} := \sigma_i(H_j)$ for all $j \le t$.

By definition of the integers d_1, d_2, we can partition the roots $\{\alpha^{(1)}, \dots, \alpha^{(d)}\}$ of $\mathfrak{m}$ according to their first coordinate, into d_1 classes $C_1, \dots, C_{d_1}$ of cardinality d_2 each: for $\kappa \le d_1$, all $\alpha^{(i)}$ in C_κ have the same first coordinate, say ζ_κ, and the ζ_κ's are pairwise distinct. Remark that $\zeta_1, \dots, \zeta_{d_1}$ are the roots of T_1.

Fix $\kappa \le d_1$ and take i such that $\alpha^{(i)}$ is in C_κ. Because K_i is primary at α, Lazard's structure theorem on bivariate lexicographic Gröbner bases [24] implies that for $j = 1, \dots, t$, $H_{i,j} = (\xi_1 - \zeta_\kappa)^{\mu_j} \tilde{H}_{i,j}$, for some polynomial $\tilde{H}_{i,j} \in \mathbb{L}[\xi_1, \xi_2]$, monic of degree ν_j in ξ_2, and of degree less than $\mu_1 - \mu_j$ in ξ_1.

For $1 \le \kappa \le d_1$ and $1 \le j \le t$, let us then define $\tilde{G}_{\kappa,j} := \prod_i \tilde{H}_{i,j}$, where the product is taken over all i such that $\alpha^{(i)} \in C_\kappa$. This is a polynomial in $\mathbb{L}[\xi_1, \xi_2]$, with leading term ${\xi_2}^{d_2 \nu_j}$. Finally, let $\tilde{G}_1 := 1$, and for $2 \le j \le t$ let $\tilde{G}_j$ be the unique polynomial in $\mathbb{L}[\xi_1, \xi_2]$ of degree less than $d_1(\mu_1 - \mu_j)$ in ξ_1 such that $\tilde{G}_j \bmod (\xi_1 - \zeta_\kappa)^{\mu_1 - \mu_j} = \tilde{G}_{\kappa,j}$ holds for all $\kappa \le d_1$. We claim that $(G_1, \dots, G_t)$, with $G_j := {T_1}^{\mu_j} \tilde{G}_j$ for all j, is a Gröbner basis of $I \cdot \mathbb{L}[\xi_1, \xi_2]$, minimal but not necessarily reduced.

4.1.5. To establish this claim, we first prove that $I \cdot \mathbb{L}[\xi_1, \xi_2] = \langle G_1, \dots, G_t \rangle$ in $\mathbb{L}[\xi_1, \xi_2]$. The first step is to determine the common zeros of $G_1, \dots, G_t$. Since $G_1 = {T_1}^{\mu_1}$, the ξ_1-coordinates of the solutions are the roots $\{\zeta_1, \dots, \zeta_{d_1}\}$ of T_1. Fix $\kappa \le d_1$, and let (ζ_κ, η) be a root of $G_1, \dots, G_t$. In particular, $G_t(\zeta_\kappa, \eta) = \tilde{G}_t(\zeta_\kappa, \eta) = 0$. This implies that $\tilde{G}_{\kappa,t}(\zeta_\kappa, \eta) = 0$, so there exists $i \le d$ such that $(\zeta_\kappa, \eta) = \alpha^{(i)}$. Conversely, any $\alpha^{(i)}$ cancels $G_1, \dots, G_t$, so that the zero-sets of $G_1, \dots, G_t$ and $I \cdot \mathbb{L}[\xi_1, \xi_2]$ are equal. Next, we determine the primary component Q_i of $\langle G_1, \dots, G_t \rangle$ at a given $\alpha^{(i)}$.

Take such an index i, and assume that $\alpha^{(i)}$ is in C_κ, for some $\kappa \le d_1$ (so the first coordinate of $\alpha^{(i)}$ is ζ_κ). Take D large enough, so that $D \ge \mu_1$ and $(\xi_1 - \zeta_\kappa)^D$ belongs to Q_i; hence Q_i is also the primary component of the ideal $\langle G_1, \dots, G_t, (\xi_1 - \zeta_\kappa)^D \rangle$ at $\alpha^{(i)}$. This ideal is generated by the polynomials $(\xi_1 - \zeta_\kappa)^{\mu_1}$ and $(\xi_1 - \zeta_\kappa)^{\mu_j} \tilde{G}_j$, for $2 \le j \le t$. For such j, since $\tilde{G}_j \bmod (\xi_1 - \zeta_\kappa)^{\mu_1 - \mu_j} = \tilde{G}_{\kappa,j}$, we get that $(\xi_1 - \zeta_\kappa)^{\mu_j} \tilde{G}_j \bmod (\xi_1 - \zeta_\kappa)^{\mu_1} = (\xi_1 - \zeta_\kappa)^{\mu_j} \tilde{G}_{\kappa,j}$. As a result, the ideal above also admits the generators $(\xi_1 - \zeta_\kappa)^{\mu_1}, (\xi_1 - \zeta_\kappa)^{\mu_2} \tilde{G}_{\kappa,2}, \dots, \tilde{G}_{\kappa,t}$. Now, recall that $\tilde{G}_{\kappa,j} = \prod_\iota \tilde{H}_{\iota,j}$, where the product is taken over all ι such that $\alpha^{(\iota)}$ is in C_κ. For $\iota \ne i$, $\tilde{H}_{\iota,j}$ does not vanish at $\alpha^{(i)}$ [24, Th. 2.(i)], so it is invertible locally at $\alpha^{(i)}$. It follows that the primary component of G at $\alpha^{(i)}$ is generated by $(\xi_1 - \zeta_\kappa)^{\mu_1}, (\xi_1 - \zeta_\kappa)^{\mu_2} \tilde{H}_{i,2}, \dots, \tilde{H}_{i,t}$, that is, $H_{i,1}, \dots, H_{i,t}$. This is precisely the ideal K_i.

To summarize, $\langle G_1, \dots, G_t \rangle$ and $I \cdot \mathbb{L}[\xi_1, \xi_2]$ have the same primary components $K_1, \dots, K_d$, so these ideals coincide. It remains to prove that $(G_1, \dots, G_t)$ is a Gröbner basis of $I \cdot \mathbb{L}[\xi_1, \xi_2]$. The shape of the leading terms of $G_1, \dots, G_t$ implies that number of monomials reduced with respect to these polynomials is $d \deg(J) = d\mu$. Now, since all its primary components K_i have degree $\mu = \deg(J)$, the ideal $I \cdot \mathbb{L}[\xi_1, \xi_2] = \langle G_1, \dots, G_t \rangle$ has degree $d\mu$ as well. As a result, $G_1, \dots, G_t$ form a Gröbner basis (since otherwise, applying the Buchberger algorithm to them would yield fewer reduced monomials, a contradiction).

The polynomials $G_1, \dots, G_t$ are a Gröbner basis, *minimal*, as can be seen from their leading terms, but not reduced; we let $R_1, \dots, R_t$ be the corresponding reduced minimal Gröbner basis. For all j, ${T_1}^{\mu_j}$ divides G_j, and we obtain R_j by reducing G_j by multiples of ${T_1}^{\mu_j}$, so that each R_j is a multiple of ${T_1}^{\mu_j}$ as well. In addition, the leading terms of G_j and R_j are the same. Hence, our proposition is proved.

4.1.6. As a corollary, the following proposition and its proof extend [22, Lemma 9] to bivariate contexts. We will still use the names *untangling* and *tangling* for $\pi_{\mathfrak{m},J'}$ as defined below and its inverse.

Proposition 4.2. *Assuming I is a zero-dimensional ideal in $\mathbb{F}[x_1, x_2]$ with a perfect field $\mathbb{F}$ of characteristic greather than $\deg(I)$ and $\mathfrak{m}$ is a maximal ideal; there exists an $\mathbb{F}$-algebra isomorphism*

$$\pi_{\mathfrak{m},J'} : \mathbb{F}[x_1, x_2]/I \to \mathbb{K}[\xi_1, \xi_2]/J' \qquad (1)$$

given by $(x_1, x_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$, where α_1, α_2 are respectivly the residue classes of y_1, y_2 in $\mathbb{K} := \mathbb{F}[y_1, y_2]/\tilde{\mathfrak{m}}$ and $J' \subset \mathbb{K}[\xi_1, \xi_2]$ the ideal obtained by applying the translating a zero-dimensional primary ideal in $\alpha := (\alpha_1, \alpha_2)$ to (0,0).

Proof. We prove that the embedding $\Phi : \mathbb{F}[x_1, x_2] \to \mathbb{K}[\xi_1, \xi_2]$ given by $(x_1, x_2) \mapsto (\xi_1, \xi_2)$ induces an isomorphism of $\mathbb{F}$-algebras $\mathbb{F}[x_1, x_2]/I \to \mathbb{K}[\xi_1, \xi_2]/J$. From this, applying the change of variables $(\xi_1, \xi_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$ gives the result.

Since $\Phi(I)$ is contained in J, the embedding Φ induces an homomorphism $\phi : \mathbb{F}[x_1, x_2]/I \to \mathbb{K}[\xi_1, \xi_2]/J$. By the previous proposition, both sides have dimension $d\mu$ over $\mathbb{F}$, so it is enough to prove that ϕ is injective. But this amounts to verifying that $\Phi^{-1}(J) = I$, which is true by definition. □

4.2 Untangling for monomial ideals

4.2.1. In this section, we give an algorithm for the mapping $\pi_{\mathfrak{m},J'}$ of Proposition 4.2 under a simplifying assumption. To state it, recall that J' is maximal at $(0,0) \in \mathbb{K}^2$. Then, our assumption is

H$_3$**.** J' is a *monomial* ideal.

In view of the shape of the leading terms given in **4.1.3** for the ideal J, we deduce that $J' = \langle \xi_1^{\mu_1}, \xi_1^{\mu_2} \xi_2^{\nu_2}, \dots, \xi_2^{\nu_t} \rangle$. In the rest of this subsection, $\mathcal{B}$ is the monomial basis of $\mathbb{F}[x_1, x_2]/I$ induced by the Gröbner basis exhibited in Proposition 4.1 and $\mathcal{B}'$ is the monomial basis of $\mathbb{K}[\xi_1, \xi_2]/J'$. Then, the inputs of the algorithms in this subsection are in $\mathrm{Span}_{\mathbb{F}}\mathcal{B} := \oplus_{b \in \mathcal{B}} \mathbb{F} b$, and the outputs in $\mathrm{Span}_{\mathbb{K}}\mathcal{B}' := \oplus_{b' \in \mathcal{B}'} \mathbb{K} b'$. This being said, our result is the following.

Proposition 4.3. *Under* $\mathbf{H}_2$ *and* $\mathbf{H}_3$*, given* F *in* $\mathbb{F}[x_1, x_2]/I$ *one can compute* $\pi_{\mathfrak{m},J'}(F)$ *using either* $O(\mathsf{M}(dn))$ *or* $O(\mathsf{M}(\mu n)\log(\mu))$ *operations in* $\mathbb{F}$*, and in particular in* $O(\mathsf{M}(n^{1.5})\log(n))$ *operations.*

We prove the first two bounds in **4.2.2** and **4.2.3** respectively. The last statement readily follows, since $n = d\mu$ (Proposition 4.1).

4.2.2. We start with an efficient algorithm for those cases where $d = [\mathbb{K} : \mathbb{F}]$ is small. The idea is simple: as in the univariate case, the untangling mapping $\pi_{\mathfrak{m},J'}$ can be rephrased in terms of Taylor expansion. Explicitly, for F in $\mathbb{F}[x_1, x_2]/I$, $\pi_{\mathfrak{m},J'}(F)$ is simply

$$F(\xi_1 + \alpha_1, \xi_2 + \alpha_2) \bmod \langle \xi_1^{\mu_1}, \xi_1^{\mu_2}\xi_2^{\nu_2}, \dots, \xi_2^{\nu_t} \rangle.$$

We compute $F(\xi_1 + \alpha_1, \xi_2 + \alpha_2)$, proceeding one variable at a time.

Step 1. Compute $F^* := F(\xi_1 + \alpha_1, \xi_2) \in \mathbb{K}[\xi_1, \xi_2]$. Because $2, \dots, n$ are units in $\mathbb{F}$, given a univariate polynomial P of degree $t \le n$ in $\mathbb{K}[\xi_1]$ one can compute $P(\xi_1 + \alpha_1)$ in $O(\mathsf{M}(t))$ operations $(+, \times)$ in $\mathbb{K}$ (see [1]). Using Kronecker substitution [17, Chapter 8.4], this translates to $O(\mathsf{M}(dt))$ operations in $\mathbb{F}$ (we will systematically use such techniques, see e.g. Lemma 2.2 in [18] for details). Computing F^* is done by applying this procedure coefficient-wise with respect to ξ_2; in particular, all ξ_1-degrees involved are at most n, and add up to n. The super-linearity of M implies that this takes a total of $O(\mathsf{M}(dn))$ operations in $\mathbb{F}$.

Step 2. Compute $F^*(\xi_1, \xi_2 + \alpha_2) = F(\xi_1 + \alpha_1, \xi_2 + \alpha_2)$. This is done in the same manner, applying the translation with respect to ξ_2 instead; the runtime is still $O(\mathsf{M}(dn))$ operations in $\mathbb{F}$.

Step 3. Since F is in $\mathrm{Span}_{\mathbb{F}}\mathcal{B}$, and $\mathcal{B}$ is stable by division, $F(\xi_1 + \alpha_1, \xi_2 + \alpha_2)$ are in $\mathrm{Span}_{\mathbb{K}}\mathcal{B} := \oplus_{b \in \mathcal{B}} \mathbb{K} b$. By Proposition 4.1, all monomials in $\mathcal{B}'$ are in $\mathcal{B}$, so we can obtain $\pi_{\mathfrak{m},J'}(F)$ by discarding from $F(\xi_1 + \alpha_1, \xi_2 + \alpha_2)$ all monomials not in $\mathcal{B}'$.

Overall, the runtime is $O(\mathsf{M}(dn))$ operations in $\mathbb{F}$. For small d, when the multiplicity μ is large, this is close to being linear in $n = \deg(I)$.

4.2.3. Next we give an another solution, which will perform well in cases where the multiplicity $\mu = \deg(J')$ is small.

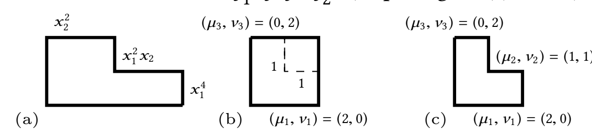
Again the idea is simple: given F in $\mathrm{Span}_{\mathbb{F}}\mathcal{B}$, compute $F(\xi_1 + \alpha_1, \xi_2 + \alpha_2) \bmod \langle {\xi_1}^{\mu_1}, {\xi_2}^{\nu_t} \rangle$, and again discard unwanted terms (this is correct, since all coefficients of $\pi_{\mathfrak{m},J'}(F)$ are among those we compute). As in the previous paragraph, this is done one variable at a time; in the following, recall that $\mathfrak{m} = \langle T_1(x_1), T_2(x_1, x_2) \rangle$, with $\deg(T_1, x_1) = d_1$ and $\deg(T_2, x_2) = d_2$, so that $d_1 d_2 = d = \deg(\mathfrak{m})$. Also, we let $\mathbb{K}'$ be the subfield $\mathbb{F}[y_1]/\langle T_1(y_1) \rangle$ of $\mathbb{K}$, so that $\mathbb{K} = \mathbb{K}'[y_2]/\langle T_2(\alpha_1, y_2) \rangle$; we have $[\mathbb{K}' : \mathbb{F}] = d_1$ and $[\mathbb{K} : \mathbb{K}'] = d_2$.

Step 1. By Proposition 4.1, we can write $F = \sum_{0 \le i < d_2 \nu_t} F_i(x_1) x_2^i$, with all F_i's of degree at most $d_1 \mu_1$. Compute all $F_i^* := \pi_{T_1, \mu_1}(F_i) \in \mathbb{K}'[\xi_1]/\langle {\xi_1}^{\mu_1} \rangle$, so as to obtain $G := \sum_{0 \le i < d_2 \nu_t} F_i^* x_2^i$. The cost of this step is $O(d_2 \nu_t \mathsf{M}(d_1 \mu_1) \log(\mu_1))$ operations in $\mathbb{F}$. Since $\nu_t \mu_1 \le \mu^2$ and $d_1 d_2 \mu = d\mu = n$, with $n = \deg(I)$, this is $O(\mathsf{M}(\mu n)\log(\mu))$.

Step 2. Rewrite G as $G = \sum_{i < \mu_1} G_i(x_2) \xi_1^i$, with all G_i's in $\mathbb{K}'[x_2]$ of degree at most $d_2 \nu_t$. Compute all $G_i^* := \pi_{T_2, \nu_t}(G_i) \in \mathbb{K}[\xi_2]/\langle {\xi_2}^{\nu_t} \rangle$.

To compute the G_i^*'s, we apply the univariate untangling algorithm with coefficients in $\mathbb{K}'$ instead of $\mathbb{F}$. The runtime of this second step is $O(\mu_1 \mathsf{M}(d_2 \nu_t)\log(\nu_t))$ operations $(+, \times)$ in $\mathbb{K}'$, which becomes $O(\mu_1 \mathsf{M}(d_1 d_2 \nu_t)\log(\nu_t))$ operations in $\mathbb{F}$, once we use Kronecker substitution to do arithmetic in $\mathbb{K}'$. As for the first step, this is $O(\mathsf{M}(\mu n)\log(\mu))$ operations in $\mathbb{F}$.

Step 3. At this stage, we have $\sum_{i < d_2 \nu_t} G_i^* {\xi_1}^i \in \mathbb{K}[\xi_2]/\langle {\xi_1}^{\mu_1}, {\xi_2}^{\nu_t} \rangle = F(\xi_1 + \alpha_1, \xi_2 + \alpha_2) \bmod \langle {\xi_1}^{\mu_1}, {\xi_2}^{\nu_t} \rangle$. Discard all monomials lying in J' and return the result – this involves no arithmetic operation. On our example, the untangling algorithm would pass from an ideal in x_1, x_2 (figure (a) below) to the monomial ideal $\langle \xi_1^2, \xi_2^2 \rangle$ (step 2, figure (b) below) then the monomial $\xi_1 \xi_2$ would be discarded to get a result defined modulo $J' = \langle \xi_1^2, \xi_1 \xi_2, \xi_2^2 \rangle$ (step 3, figure (c) below).



4.3 Recursive tangling for monomial ideals

The ideas used to perform univariate tangling, that is, invert $\pi_{T,\mu}$, carry over to bivariate situations. In this section, we discuss the first of them, namely, a bivariate version of van der Hoeven and Lecerf's recursive algorithm. We still work under assumption $\mathbf{H}_3$ that J' is a monomial ideal. As before, $\mathcal{B}$ is the monomial basis of $\mathbb{F}[x_1, x_2]/I$ induced by the Gröbner basis exhibited in Proposition 4.1.

Proposition 4.4. *Under* $\mathbf{H}_2$ *and* $\mathbf{H}_3$*, given* G *in* $\mathbb{K}[\xi_1, \xi_2]/J'$ *one can compute* ${\pi_{\mathfrak{m},J'}}^{-1}(G)$ *using either* $O(\mathsf{M}(dn)\log(n) + \mathsf{M}(n)\log(n)^2)$*, or* $O(\mathsf{M}(\mu n)\log(n)^2)$ *operations in* $\mathbb{F}$*. In particular, this can be done in* $O(\mathsf{M}(n^{1.5})\log(n)^2)$ *operations.*

As in [22], our procedure is recursive; the recursion here is based on the integer μ_1. Given G in $\mathbb{K}[\xi_1, \xi_2]/J'$, we explain how to find F in $\mathbb{F}[x_1, x_2]/I$ such that $\pi_{\mathfrak{m},J'}(F) = G$, starting from the case $\mu_1 = 1$.

4.3.1. If $\mu_1 = 1$, the ideal J' is of the form $\langle \xi_1, {\xi_2}^{\nu_2} \rangle$, and $\pi_{\mathfrak{m},J'}$ maps $F(x_1, x_2)$ to $G := F(\alpha_1, \xi_2 + \alpha_2) \bmod {\xi_2}^{\nu_2}$. In this case, note that the degree n of I is simply $d_1 d_2 \nu_2$.

Step 1. Apply our univariate tangling algorithm to G in the variable x_2, to compute $F(\alpha_1, x_2) := \pi_{T_2,\nu_2}^{-1}(G) \in \mathbb{K}'[x_2]/\langle T_2^{\mu_2} \rangle$; we work over the field $\mathbb{K}' = \mathbb{F}[y_1]/\langle T_1(y_1) \rangle$ instead of $\mathbb{F}$. This takes $O(\mathsf{M}(d_2 \nu_2)\log(\nu_2) + \mathsf{M}(d_2)\log(d_2))$ operations $(+, \times)$ in $\mathbb{K}'$, together with $O(d_2)$ inversions in $\mathbb{K}'$. Using Kronecker substitution for multiplications, this results in a total of $O(\mathsf{M}(d_1 d_2 \nu_2)\log(\nu_2) + \mathsf{M}(d_1 d_2)\log(d_1 d_2))$ operations in $\mathbb{F}$. We will use the simplified upper bound $O(\mathsf{M}(d_1 d_2 \nu_2)\log(d_1 d_2 \nu_2)) = O(\mathsf{M}(n)\log(n))$.

Step 2. The polynomial F has degree less than d_1 in x_1 and $d_2 \nu_2$ in x_2; for such F's, knowing $F(\alpha_1, x_2) \in \mathbb{K}'[x_2]/\langle T_2^{\mu_2} \rangle$ is equivalent to knowing $F(x_1, x_2)$ in $\mathbb{F}[x_1, x_2]$. Thus, we are done.

4.3.2. Assume now that $\mu_1 > 1$, let G be in $\mathbb{K}[\xi_1, \xi_2]/J'$ and let $\bar{\mu} := \lceil \mu_1/2 \rceil$. The following steps follow closely Algorithm 9 in [22]. For the cost analysis, we let $S(\mathfrak{m}, J')$ be the cost of applying $\pi_{\mathfrak{m},J'}$ (see Proposition 4.4) and $T(\mathfrak{m}, J')$ be the cost of the recursive algorithm for ${\pi_{\mathfrak{m},J'}}^{-1}$.

Step 1. Let $\bar{G} := G \bmod {\xi_1}^{\bar{\mu}}$, and compute recursively $\bar{F} := {\pi_{\mathfrak{m},J_0'}}^{-1}(\bar{G})$, with $J_0' := J' + \langle {\xi_1}^{\bar{\mu}} \rangle$. This costs $T(\mathfrak{m}, J_0')$.

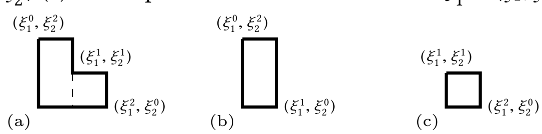
Step 2. Compute $H := (G - \pi_{\mathfrak{m},J'}(\bar{F}))$ div ${\xi_1}^{\bar{\mu}}$, where the div operator maps ${\xi_1}^i$ to 0 for $i < \bar{\mu}$ and to ${\xi_1}^{i-\bar{\mu}}$ otherwise. This costs $S(\mathfrak{m}, J')$.

Step 3. Define $W := \xi_1/\pi_{\mathfrak{m},J'}(T_1) \in \mathbb{K}[\xi_1,\xi_2]/\langle {\xi_1}^{\mu_1}, {\xi_2}^{\mu_2}\rangle$. Because $T_1(\alpha_1) = 0$ and $T_1'(\alpha_1) \neq 0$ (by the separability assumption), W is well-defined. This costs $S(\mathfrak{m}, J')$ for $\pi_{\mathfrak{m},J'}(T_1)$ and $O(\mathsf{M}(d_1\mu_1))$ for inversion (since it involves ξ_1 only), which is $O(\mathsf{M}(n))$.

Step 4. Compute recursively $\bar{E} := {\pi_{\mathfrak{m},J_1'}}^{-1}(W^{\bar{\mu}} H \bmod J_1')$, where J_1' is the colon ideal $J' : {\xi_1}^{\bar{\mu}}$. Since W depends only on ξ_1, a multiplication by W, or one of its powers, is done coefficient-wise in ξ_2, for $O(\mathsf{M}(n))$ operations in $\mathbb{F}$. Thus, the cost to compute $W^{\bar{\mu}} H \bmod J_1'$ is $O(\mathsf{M}(n)\log(n))$; to this, we add $T(\mathfrak{m}, J_1')$.

Step 5. Return $F := \bar{F} + {T_1}^{\bar{\mu}} \bar{E}$. The product ${T_1}^{\bar{\mu}} \bar{E}$ requires no reduction, since all its terms are in $\mathcal{B}$. Proceeding coefficient-wise with respect to x_2, and using super-additivity, it costs $O(\mathsf{M}(n))$.

On our example, we have $J' = \langle \xi_1^2, \xi_1\xi_2, \xi_2^2\rangle$ (a), Step 1 uses $J_0' = \langle \xi_1, \xi_2^2\rangle$ (b) and Steps 2-5 work on the colon ideal $J_1' = \langle \xi_1, \xi_2\rangle$ (c).



Let us justify that this algorithm is correct, by computing $\pi_{\mathfrak{m},J'}(F)$, which is equal to $\pi_{\mathfrak{m},J'}(\bar{F}) + \pi_{\mathfrak{m},J'}(T_1)^{\bar{\mu}} \pi_{\mathfrak{m},J'}(\bar{E}) \bmod J'$. Note first that $\pi_{\mathfrak{m},J'}(\bar{F}) \bmod {\xi_1}^{\bar{\mu}} = G \bmod {\xi_1}^{\bar{\mu}}$. Equivalently, $\pi_{\mathfrak{m},J'}(\bar{F}) = G \bmod {\xi_1}^{\bar{\mu}} + {\xi_1}^{\bar{\mu}}(\pi_{\mathfrak{m},J'}(\bar{F}) \text{ div } {\xi_1}^{\bar{\mu}})$. Using the definition of H, this is also $G \bmod {\xi_1}^{\bar{\mu}} + {\xi_1}^{\bar{\mu}}(G \text{ div } {\xi_1}^{\bar{\mu}} - H)$, that is, $G - {\xi_1}^{\bar{\mu}} H$. On the other hand, by definition of $\bar{E}$, we have

$$\pi_{\mathfrak{m},J'}(\bar{E}) = \pi_{\mathfrak{m},J'}({\pi_{\mathfrak{m},J_1'}}^{-1}(W^{\bar{\mu}} H \bmod J_1')),$$

so that $\pi_{\mathfrak{m},J'}(\bar{E}) \bmod J_1' = W^{\bar{\mu}} H \bmod J_1'$. Now, $\pi_{\mathfrak{m},J'}(T_1)$ is a multiple of ξ_1, so $\pi_{\mathfrak{m},J'}(T_1)^{\bar{\mu}}$ is a multiple of ${\xi_1}^{\bar{\mu}}$. Since ${\xi_1}^{\bar{\mu}} J_1'$ is in J', we deduce that $\pi_{\mathfrak{m},J'}(T_1)^{\bar{\mu}} \pi_{\mathfrak{m},J'}(\bar{E}) \bmod J'$ is equal to $\pi_{\mathfrak{m},J'}(T_1)^{\bar{\mu}} W^{\bar{\mu}} H \bmod J'$, and thus to ${\xi_1}^{\bar{\mu}} H$. Adding the two intermediate results so far, we deduce that $\pi_{\mathfrak{m},J'}(F) = G$, as claimed.

Finally, we do the cost analysis. The runtime $T(\mathfrak{m}, J')$ satisfies the recurrence relation

$$T(\mathfrak{m}, J') = T(\mathfrak{m}, J_0') + T(\mathfrak{m}, J_1') + O(S(\mathfrak{m}, J') + \mathsf{M}(n)\log(n)).$$

Using **4.3.1** and the super-linearity of M, we see that the total cost at the leaves is $O(\mathsf{M}(n)\log(n))$. Without loss of generality, we can assume that $S(\mathfrak{m}, J')$ is super-linear, in the sense that $S(\mathfrak{m}, J_0') + S(\mathfrak{m}, J_1') \leq S(\mathfrak{m}, J')$ holds at every level of the recursion. Since the recursion has depth $O(\log(n))$, we get that $T(\mathfrak{m}, J')$ is in $O(S(\mathfrak{m}, J')\log(n) + \mathsf{M}(n)\log(n)^2)$.

4.4 Tangling for monomial ideals using duality

We finally present a bivariate analogue of the algorithm introduced in Section 3. Since the runtimes obtained are in general worse than those in the previous subsection, we only sketch the construction.

All notation being as before, let G be in $\mathbb{K}[\xi_1,\xi_2]/J'$, and let $F \in \mathbb{F}[x_1,x_2]/I$ be such that $\pi_{\mathfrak{m},J'}(F) = G$. Following ideas from [30], we now use several linear forms. Thus, let $\ell_1, \dots, \ell_\gamma$ be module generators of $(\mathbb{K}[\xi_1,\xi_2]/J')^*$, where the * means that we look at the dual of $\mathbb{K}[\xi_1,\xi_2]/J'$ as an $\mathbb{F}$-vector space. Define $\ell_1' := G \cdot \ell_1, \dots, \ell_\gamma' := G \cdot \ell_\gamma$, as well as

$$L_1 := \pi_{\mathfrak{m},J'}^{\perp}(\ell_1), \dots, L_\gamma := \pi_{\mathfrak{m},J'}^{\perp}(\ell_\gamma)$$
$$L_1' := \pi_{\mathfrak{m},J'}^{\perp}(\ell_1'), \dots, L_\gamma' := \pi_{\mathfrak{m},J'}^{\perp}(\ell_\gamma')$$

in $(\mathbb{F}[x_1,x_2]/I)^*$. As in one variable, for $i = 1, \dots, \gamma$ the relation $\pi_{\mathfrak{m},J'}(F) \cdot \ell_i = \ell_i'$ implies that $F \cdot L_i = L_i'$.

The first question is to determine suitable $\ell_1, \dots, \ell_\gamma$. Consider generators $\xi_1^{\mu_1}\xi_2^{\nu_1}, \dots, \xi_1^{\mu_t}\xi_2^{\nu_t}$ of J', with the μ_i's decreasing and ν_i's increasing as before. For $i = 1, \dots, t-1$, define ℓ_i by $\ell_i(\alpha_1^{d_1-1}\alpha_2^{d_2-1}{\xi_1}^{\mu_i-1}{\xi_2}^{\nu_{i+1}-1}) = 1$, all other $\ell_i(\alpha_1^{e_1}\alpha_2^{e_2}{\xi_1}^{r_1}{\xi_i}^{r_2})$ being set to zero. Then, following e.g. [15, Section 21.1], one verifies that these linear forms are module generators of $(\mathbb{K}[\xi_1,\xi_2]/J')^*$.

As in the univariate case, we can compute all L_i and L_i' by transposing the untangling algorithm, incurring $O(t)$ times the cost reported in Proposition 4.4. Then, it remains to solve all equations $F \cdot L_i = L_i'$, $i = 1, \dots, t-1$ (this system is not square, unless $t = 2$). We are not aware of a quasi-linear time algorithm to solve such systems. The matrix of an equation such as $F \cdot L_i = L_i'$ is sometimes called *multi-Hankel* [4]. It can be solved using structured linear algebra techniques [4] (Here, we have several such systems to solve at once; this can be dealt with as in [11]). As in [4], using the results from [6] on structured linear system solving, we can find F in Monte Carlo time $O((st)^{\omega-1}\mathsf{M}(tn)\log(tn))$, with $s := \min(\mu_1, \nu_t)$, where ω is the exponent of linear algebra (the best value to date is $\omega \leq 2.38$ [12, 25]). Thus, unless both s and t are small, the overhead induced by the linear algebra phase may make this solution inferior to the one in the previous subsection.

4.5 Application

To conclude, we describe a direct application of our results, to the complexity of multiplication and inverse in $\mathbb{A} := \mathbb{F}[x_1,x_2]/I$: under assumptions $\mathbf{H}_2$ and $\mathbf{H}_3$, both can be done in the time reported in Proposition 4.4, to which we add $O(\mathsf{M}(n)\log(n)^3)$ in the case of inversion. Even though the algorithms are not quasi-linear time in the worst case, to our knowledge no previous non-trivial algorithm was known for such operations.

The algorithms are simple: untangle the input, do the multiplication, resp. inversion, in $\mathbb{A}' := \mathbb{K}[\xi_1,\xi_2]/J'$, and tangle the result. The cost of tangling dominates that of untangling. The appendix below discusses the cost of arithmetic in $\mathbb{A}'$: multiplication and inverse take respectively $O(\mathsf{M}(\mu)\log(\mu))$ and $O(\mathsf{M}(\mu)\log(\mu)^2)$ operations $(+,-,\times)$ in $\mathbb{K}$, plus one inverse in $\mathbb{K}$ for the latter. Using Kronecker substitution, the runtimes become $O(\mathsf{M}(n)\log(n))$ and $O(\mathsf{M}(n)\log(n)^2)$ operations in $\mathbb{K}$, with $n = \deg(I)$; this is thus negligible in front of the cost for tangling.

Appendix: Bivariate power series arithmetic

We prove that for a field $\mathbb{F}$ and zero-dimensional monomial ideal $I \subset \mathbb{F}[x_1,x_2]$, multiplication and inversion in $\mathbb{F}[x_1,x_2]/I$ can be done in softly linear time in $\delta := \deg(I)$, starting with multiplication.

For an ideal such as $I = \langle x_1^{\mu}, x_2^{\nu}\rangle$, the claim is clear. Indeed, to multiply elements F and G of $\mathbb{F}[x_1,x_2]/I$ we multiply them as bivariate polynomials and discard unwanted terms. Bivariate multiplication in partial degrees less than μ, resp. ν, can be done by Kronecker substitution in time $O(\mathsf{M}(\mu\nu)) = O(\mathsf{M}(\delta))$, which is softly linear in δ, as claimed. However, this direct approach does not perform well for cases such as $I = \langle x_1^{\mu}, x_1 x_2, x_2^{\nu} \rangle$: in this case, for F and G reduced modulo I, the product FG as polynomials has $\mu\nu$ terms, but $\delta = \mu + \nu - 1$. The following result shows that in general, we can obtain a cost almost as good as in the first case, up to a logarithmic factor. Whether this extra factor can be removed is unclear to us. In the rest of the annexe, we write $I = \langle x_1^{\mu_1} x_2^{\nu_1}, x_1^{\mu_2} x_2^{\nu_2}, \dots, x_1^{\mu_t} x_2^{\nu_t} \rangle$, with μ_i's decreasing, ν_i's increasing and $\nu_1 = \mu_t = 0$.

Proposition 4.5. *Let I be a zero-dimensional monomial ideal in $\mathbb{F}[x_1, x_2]$ of degree δ. Given F, G reduced modulo I, one can compute $FG \bmod I$ in $O(\mathsf{M}(\delta)\log(\delta))$ operations $(+,-,\times)$ in $\mathbb{F}$.*

A.1. We start by giving an algorithm of complexity $O(t\mathsf{M}(\delta))$ for multiplication modulo I. Let F and G be two polynomials reduced modulo I. To compute $H := FG \bmod I$ it suffices to compute $H_i := FG \bmod \langle x_1^{\mu_i}, x_2^{\nu_{i+1}} \rangle$ for $i = 1, \dots, t-1$; all monomials in H appear in one of the H_i's (some of them in several H_i's). We saw that multiplication modulo $\langle x_1^{\mu_i}, x_2^{\nu_{i+1}} \rangle$ takes $O(\mathsf{M}(\mu_i \nu_{i+1}))$ operations in $\mathbb{F}$, which is $O(\mathsf{M}(\delta))$, so the total cost is $O(t\mathsf{M}(\delta))$.

A.2. In the general case, define $i_1 := 1$. We let $i_2 \le t$ be the smallest index greater than i_1 and such that $\mu_{i_2} < \mu_{i_1}/2$, and iterate the process to define a sequence $i_1 = 1 < i_2 < \cdots < i_s = t$. The ideal I' is then defined by the monomials $x_1^{\mu_{i_1}} x_2^{\nu_{i_1}}, \dots, x_1^{\mu_{i_s}} x_2^{\nu_{i_s}}$. By construction, I contains I'; hence, to compute a product modulo I, we may compute it modulo I' and discard unwanted terms.

Multiplication modulo I' is done using the algorithm of **A.1**, in time $O(s\mathsf{M}(\delta'))$, with $\delta' := \deg(I')$. Hence, we need to estimate the degree δ' of I', as well as its number of generators s.

The degree δ of I can be written as $\sum_{r=1}^{s-1} \sum_{i=i_r}^{i_{r+1}-1} \mu_i(\nu_{i+1} - \nu_i)$; this is simply counting the number of standard monomials along the rows. For a given r, all indices i in the inner sum are such that $\mu_i \ge \mu_{i_r}/2$, so the sum is at least $1/2 \sum_{r=1}^{s-1} \mu_{i_r}(\nu_{i_{r+1}} - \nu_{i_r})$, which is the degree of I'. Hence, $\delta \ge 1/2\delta'$, that is, $\delta' \le 2\delta$. To estimate the number s, the inequalities $\mu_{i_{r+1}} < \mu_{i_r}/2$ for all $r \le s$ imply that $\mu_{i_{s-1}} < \mu_1/2^s$. We deduce that $2^s \le \mu_1/\mu_{i_{s-1}} \le \mu_1$ (since $\mu_{i_{s-1}} \ge 1$), which itself is at most δ. Thus, $s \in O(\log(\delta))$. Overall, the cost of multiplication modulo I', and thus modulo I, is $O(\mathsf{M}(\delta)\log(\delta))$.

Corollary 4.6. *For I as in the previous proposition and F reduced modulo I, with $F(0,0) \neq 0$, $1/F \bmod I$ can be computed in $O(\mathsf{M}(\delta)\log(\delta)^2)$ operations $(+,-,\times)$ in $\mathbb{F}$, and one inverse.*

A.3. We proceed by induction using Newton iteration. If $\mu_1 = 1$ then $I = \langle x_1, x_2^{\nu_2} \rangle$, so inversion modulo I is inversion in $\mathbb{F}[x_2]/\langle x_2^{\nu_2} \rangle$. It can be done in time $O(\mathsf{M}(\delta))$ using univariate Newton iteration, involving only the inversion of the constant term of the input.

Otherwise, define $\bar{\mu} := \lceil \mu_1/2 \rceil$, and let $\bar{I}$ be the ideal with generators $x_1^{\bar{\mu}}, x_1^{\mu_2} x_2^{\nu_2}, \dots, x_2^{\nu_t}$ (all monomials in this list with $\mu_i \ge \bar{\mu}$ may be discarded). Given F in $\mathbb{F}[x_1, x_2]/I$, we start by computing the inverse of $\bar{G}$ of $\bar{F} := F \bmod \bar{I}$ in $\mathbb{F}[x_1, x_2]/\bar{I}$. Since $\bar{I}^2$ is contained in I, knowing $\bar{G}$, one step of Newton iteration allows us to compute $G := 1/F \bmod I$ as $G = 2\bar{G} - \bar{G}^2 F \bmod I$. Using the previous proposition, we deduce G from $\bar{G}$ in $O(\mathsf{M}(\delta)\log(\delta))$ operations. We repeat the recursion for $O(\log(\delta))$ steps, and the degrees of the ideals we consider decrease, so the overall runtime is $O(\mathsf{M}(\delta)\log(\delta)^2)$.

References

[1] A. V. Aho, K. Steiglitz, and J. D. Ullman. 1975. Evaluating polynomials at fixed sets of points. *SIAM J. Comp.* 4, 4 (1975), 533–539.

[2] M. E. Alonso, E. Becker, M.-F. Roy, and T. Wörmann. 1996. Zeroes, multiplicities and idempotents for zerodimensional systems. In *MEGA'94*. Birkhäuser, 1–15.

[3] P. Aubry, D. Lazard, and M. Moreno Maza. 1999. On the theories of triangular sets. *J. Symb. Comput.C* 28, 1,2 (1999), 45–124.

[4] J. Berthomieu, B. Boyer, and J.-C. Faugère. 2017. Linear algebra for computing Gröbner bases of linear recursive multidimensional sequences. *J. Symb. Comput.* 83 (2017), 36–67.

[5] A. Bostan, X. Caruso, G. Christol, and P. Dumas. 2018. Fast coefficient computation for algebraic power series in positive characteristic. In *ANTS-XIII*. Mathematical Sciences Publishers.

[6] A. Bostan, C. Jeannerod, C. Mouilleron, and É. Schost. 2017. On matrices with displacement structure: generalized operators and faster algorithms. *SIAM J. Matrix Anal. Appl.* 38, 3 (2017), 733–775.

[7] A. Bostan, G. Lecerf, and É. Schost. 2003. Tellegen's principle into practice. In *ISSAC'03*. ACM, 37–44.

[8] R. P. Brent, F. G. Gustavson, and D. Y. Y. Yun. 1980. Fast solution of Toeplitz systems of equations and computation of Padé approximants. *J. of Algorithms* 1, 3 (1980), 259–295.

[9] J. Canny. 1990. Generalised characteristic polynomials. *J. Symb. Comput.* 9 (1990), 241–250.

[10] J. Canny, E. Kaltofen, and Y. Lakshman. 1989. Solving systems of non-linear polynomial equations faster. In *ISSAC'89*. ACM, 121–128.

[11] M. F. I. Chowdhury, C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. 2015. Faster algorithms for multivariate interpolation with multiplicities and simultaneous polynomial approximations. *IEEE Transactions on Information Theory* 61, 5 (2015), 2370–2387.

[12] D. Coppersmith and S. Winograd. 1990. Matrix multiplication via arithmetic progressions. *J. Symb. Comput.* 9, 3 (1990), 251–280.

[13] X. Dahan. 2017. GCD modulo a primary triangular set of dimension zero. In *ISSAC'17*. ACM, 109–116.

[14] L. De Feo and É. Schost. 2012. Fast arithmetics in Artin-Schreier towers over finite fields. *J. Symb. Comput.* 47, 7 (2012), 771–792.

[15] D. Eisenbud. 1996. *Commutative Algebra with a view toward Algebraic Geometry.* Number 150 in GTM. Springer.

[16] C. M. Fiduccia. 1985. An efficient formula for linear recurrences. *SIAM J. Comput.* 14, 1 (1985), 106–112.

[17] J. Gathen and J. Gerhard. 2013. *Modern Computer Algebra* (third ed.). Cambridge University Press, Cambridge.

[18] J. von zur Gathen and V. Shoup. 1992. Computing Frobenius maps and factoring polynomials. *Computational Complexity* 2, 3 (1992), 187–224.

[19] M. Giesbrecht. 1995. Nearly optimal algorithms for canonical matrix forms. *SIAM J. Comput.* 24, 5 (1995), 948–969.

[20] M. Giusti and J. Heintz. 1993. La détermination des points isolés et de la dimension d'une variété peut se faire en temps polynomial. In *Computational algebraic geometry and commutative algebra*, Vol. XXXIV. 216–256.

[21] J. D. Hauenstein, B. Mourrain, and A. Szanto. 2017. On deflation and multiplicity structure. *J. Symb. Comput.* 83 (2017), 228–253.

[22] J. van der Hoeven and G. Lecerf. 2017. Composition modulo powers of polynomials. In *ISSAC'17*. ACM, 445–452.

[23] M. Kaminski, D.G. Kirkpatrick, and N.H. Bshouty. 1988. Addition requirements for matrix and transposed matrix products. *J. Algorithms* 9, 3 (1988), 354–364.

[24] D. Lazard. 1985. Ideal bases and primary decomposition: case of two variables. *J. Symbolic Comput.* 1, 3 (1985), 261–270.

[25] F. Le Gall. 2014. Powers of tensors and fast matrix multiplication. In *ISSAC'14*. ACM, 296–303.

[26] R. Lebreton, E. Mehrabi, and É. Schost. 2013. On the complexity of solving bivariate systems: the case of non-singular solutions. In *ISSAC'13*. ACM, 251–258.

[27] F. S. Macaulay. 1916. *The Algebraic Theory of Modular Systems.* Cambridge University Press.

[28] M. G. Marinari, H. M. Möller, and T. Mora. 1996. On multiplicities in polynomial system solving. *Trans. Amer. Math. Soc.* 348, 8 (1996), 3283–3321.

[29] E. Mehrabi and É. Schost. 2016. A softly optimal Monte Carlo algorithm for solving bivariate polynomial systems over the integers. *J. Complexity* 34 (2016), 78–128.

[30] V. Neiger, H. Rahkooy, and É. Schost. 2017. Algorithms for zero-dimensional ideals using linear recurrent sequences. In *CASC'17*. Springer, 313–328.

[31] A. Ranum. 1911. The general term of a recurring series. *Bull. Amer. Math. Soc.* 17, 9 (1911), 457–461.

[32] F. Rouillier. 1999. Solving zero-dimensional systems through the Rational Univariate Representation. *AAECC* 9, 5 (1999), 433–461.

[33] V. Shoup. 1994. Fast construction of irreducible polynomials over finite fields. *J. Symb. Comput.* 17, 5 (1994), 371–391.

[34] V. Shoup. 1999. Efficient computation of minimal polynomials in algebraic extensions of finite fields. In *ISSAC'99*. ACM, 53–58.

[35] Joris Van Der Hoeven and Robin Larrieu. 2018. Fast reduction of bivariate polynomials with respect to sufficiently regular Gröbner bases. (2018).

[36] W. T. Wu. 1986. On zeros of algebraic equations — an application of Ritt principle. *Kexue Tongbao* 31 (1986), 1–5.