Algorithms in Intersection Theory in the Plane

by

Catherine St-Pierre

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Guilaume Moroz Université de Lorraine, CNRS, Inria, LORIA Nancy, France
Supervisor(s):	Eric Schost Professor, Dept. of Computer Science, University of Waterloo
Internal Member:	Mark Giesbrecht Professor, Dept. of Computer Science, University of Waterloo
Internal-External Member:	Jason Bell Professor, Dept. of Pure Mathematics, University of Waterloo
Other Member(s):	Rafael Oliveira Assitant professor, Dept. of Computer science, University of Waterloo

Author's Declaration

This thesis consists of material, all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Chapter 1 and 2 were solely authored by Catherine St-Pierre.

Chapter 3 was co-authored by Seung Gyu Hyun, Stephen Melczer, Éric Schost and Catherine St-Pierre and published in the proceedings of the 2019's International Symposium on Symbolic and Algebraic Computation (ISSAC'19), *cf.* S. G. Hyun, S. Melczer, É. Schost, and C. St-Pierre. Change of basis for $\mathfrak{m}$-primary ideals in one and two variables. In *ISSAC'19*, pages 227–234. ACM Press, 2019 (doi: 10.1145/3326229.332626). ACM agreement with authors grants the reproduction of the manuscript with rights in the thesis hereof. All authors share equal contributions to the manuscript. Catherine St-Pierre was the sole presenter at the Symposium.

Chapter 4 was co-authored by Éric Schost and Catherine St-Pierre and is submitted for publication, *cf.* É. Schost and C. St-Pierre. Newton iteration for lexicographic gröbner bases in two variables, 2023. preprint on arXiv:2302.03766. Both authors contributed equally to the manuscript.

Chapter 5 was co-authored by Éric Schost and Catherine St-Pierre. A short version of the manuscript was accepted for the 2023's proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC'23) and will be published in ACM Press, *cf.* É. Schost, and C. St-Pierre. p-adic algorithm for bivariate Gröbner bases. In *ISSAC'23*, pages (to be determined). ACM Press, 2023 (doi: 10.1145/3597066.3597086). ACM agreement with authors grants the reproduction of the manuscript with rights in the thesis hereof. Both authors contributed equally to the manuscript.

Chapter 6 and the back matter of the thesis were solely authored by Catherine St-Pierre.

Abstract

This thesis presents an algorithm to find the local structure of intersections of plane curves. More precisely, we address the question of describing the scheme of the quotient ring of a bivariate zero-dimensional ideal $I \subseteq \mathbb{K}[x, y]$, *i.e.* finding the points (maximal ideals of $\mathbb{K}[x, y]/I$) and describing the regular functions on those points. A natural way to address this problem is via Gröbner bases as they reduce the problem of finding the points to a problem of factorisation, and the sheaf of rings of regular functions can be studied with those bases through the division algorithm and localisation. Let $I \subseteq \mathbb{K}[x, y]$ be an ideal generated by $\mathcal{F}$, a subset of $\mathbb{A}[x, y]$ with $\mathbb{A} \hookrightarrow \mathbb{K}$ and $\mathbb{K}$ a field. We present an algorithm that features a quadratic convergence to find a Gröbner basis of I or its primary component at the origin.

We introduce an $\mathfrak{m}$-adic Newton iteration to lift the lexicographic Gröbner basis of any finite intersection of zero-dimensional primary components of I if $\mathfrak{m} \subseteq \mathbb{A}$ is a *good* maximal ideal. It relies on a structural result about the syzygies in such a basis due to Conca *&* Valla [40], from which arises an explicit map between ideals in a stratum (or Gröbner cell) and points in the associated moduli space. We also qualify what makes a maximal ideal $\mathfrak{m}$ suitable for our filtration.

When the field $\mathbb{K}$ is *large enough*, endowed with an Archimedean or ultrametric valuation, and admits a fraction reconstruction algorithm, we use this result to give a complete $\mathfrak{m}$-adic algorithm to recover $\mathcal{G}$, the Gröbner basis of I. We observe that previous results of Lazard that use Hermite normal forms to compute Gröbner bases for ideals with two generators can be generalised to a set of n generators. We use this result to obtain a bound on the height of the coefficients of $\mathcal{G}$ and to control the probability of choosing a *good* maximal ideal $\mathfrak{m} \subseteq \mathbb{A}$ to build the $\mathfrak{m}$-adic expansion of $\mathcal{G}$. Inspired by Pardue [134][61, §15.9], we also give a constructive proof to characterise a Zariski open set of $\mathrm{GL}_2(\mathbb{K})$ (with action on $\mathbb{K}[x, y]$) that changes coordinates in such a way as to ensure the initial term ideal of a zero-dimensional I becomes Borel-fixed when $|\mathbb{K}|$ is sufficiently large. This sharpens our analysis to obtain, when $\mathbb{A} = \mathbb{Z}$ or $\mathbb{A} = k[t]$, a complexity less than cubic in terms of the dimension of $\mathbb{Q}[x, y]/\langle \mathcal{G} \rangle$ and softly linear in the height of the coefficients of $\mathcal{G}$.

We adapt the resulting method and present the analysis to find the $\langle x, y \rangle$-primary component of I. We also discuss the transition towards other primary components via linear mappings, called *untangling* and *tangling*, introduced by van der Hoeven and Lecerf [91]. The two maps form one isomorphism to find points with an isomorphic local structure and, at the origin, bind them. 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 give a bound on the arithmetic complexity for certain algebras.

Acknowledgements

I thank NSERC (Graham Bell Canada Graduate Scholarship), FQRNT and the University of Waterloo for their generous support.

I thank the University of Paris-Saclay (Inria) for welcoming me and giving me the best conditions to finish the redaction.

Many thanks to Prof Willard (Groups and rings, Logic), Dr Knight (Commutative algebra), Prof Bell (Advance topic groups and ring, Representation theory), Prof McKinnon (Algebraic geometry), Prof Liu (Galois theory) and Prof Satriano (Stacks and Toric varieties), Prof Schost (Symbolic computing), Prof Lau (Randomized algorithms), Prof Nehaniv (Algebraic structure of discrete dynamical systems) and Prof Ragde (Logic) for their teaching.

I would like to thank Prof Moraru, Prof Satriano, Eric Boulter and Sean Monahan for the algebraic geometry and horospherical seminars. I learned so much from them. To help me cultivate my curiosity and answer my questions even when they did not have to: Prof Schost, Prof Bell, Prof Satriano and my dearest pure mathematic friends, a learning shared is learning not forgotten; I will never forget any of you.

Thanks to Prof Schost, Prof Satriano, Dr Delisle, Sean Monahan, Eric Boulter, Nicolas Banks, Yash Vardhan Singh and Jérémy Champagne for their comments and suggestions on sections of this thesis.

Thanks to my coauthors, I had the best time working with you and learned so much – hopefully, a first step of a lifetime collaboration.

Thanks to my committee for their suggestions on how to improve this thesis.

The greatest thanks to my advisor Éric (Prof Schost) for his guidance and help. I am proud, lucky, and glad I got the chance to be your student.

Table of Contents

List of Figures **x**

List of Tables **xi**

1 Introduction **1**

1.1 Context 1

1.2 The Results 5

1.2.1 Examples 11

2 Observations **17**

2.1 Review of the key concepts 17

2.1.1 Algebraic varieties and Affine Schemes 17

2.1.2 Primary decomposition 26

2.1.3 Gröbner bases 29

2.1.4 Connection to affine varieties 38

2.1.5 Connection with schemes 38

2.1.6 Limits 40

2.1.7 Moduli Space of Strata 45

2.2 State of the Art 50

2.2.1 Folklore of arithmetic complexity 50

2.2.2 History 51

3 Change of Basis for $\mathfrak{m}$-primary Ideals in One and Two Variables **57**

3.1 Introduction 57

3.2 Preliminaries 60

3.2.1 60

3.2.2 60

3.3 The univariate case revisited 61

3.3.1 A faster tangling algorithm 62

3.3.2 Applications 65
3.4 The bivariate case 66
3.4.1 Setup 66
3.4.2 Untangling for monomial ideals 70
3.4.3 Recursive tangling for monomial ideals 72
3.4.4 Tangling for monomial ideals using duality 74
3.4.5 An Application 75

4 Newton iteration for lexicographic Gröbner bases in two variables 78
4.1 Introduction 78
4.2 Initial segments in $\mathbb{N}^2$ 86
4.2.1 Basic definitions 86
4.2.2 Structured polynomial multiplication 89
4.3 Lexicographic Gröbner bases 92
4.3.1 The structure theorem 93
4.3.2 Conca and Valla's parametrization 94
4.4 Reduction modulo a lexicographic Gröbner basis 97
4.4.1 A paving problem 97
4.4.2 Mixed radix representation 102
4.4.3 The main algorithm 104
4.5 From Gröbner parameters to Gröbner basis 112
4.6 Computing the Gröbner parameters 116
4.6.1 Starting from a reduced basis 116
4.6.2 Polynomial equations for the Gröbner parameters 121
4.7 Newton iteration 125
4.8 Conclusion 131

5 $\mathfrak{m}$-adic algorithm for bivariate Gröbner bases 132
5.1 Introduction 132
5.1.1 Leitfaden 135
5.2 Lexicographic Gröbner bases via matrix normal forms 135
5.2.1 Detaching bases 136
5.2.2 Using Hermite normal forms 137
5.2.3 Using the Howell form 141
5.3 Coefficient size and bad reductions 146
5.3.1 Coefficient size and bad reductions for Hermite normal forms 148
5.3.2 Application to Gröbner bases 149
5.4 Applying generic changes of coordinates 150

5.4.1 Equations in general position 150
5.4.2 The initial ideal is Borel-fixed 151
5.5 Main algorithms 156
5.5.1 Overview of the algorithm and its subroutines 157
5.5.2 Algorithm over general $\mathbb{A}$ 161
5.5.3 Parameters choice 161
5.5.4 Runtime analysis 163
5.5.5 Variant 164

6 Discussion **167**
6.0.1 Note on Newton iteration and field extension 168
6.0.2 The base ring 168
6.0.3 Borel-fixed 169
6.1 A few open questions 169
6.1.1 Bound on the coefficients 170
6.1.2 Primary decomposition 170

References **172**

APPENDICES **185**

A Products **186**
A.1 Tensor product 186
A.2 Exterior product 187

B Special Matrices **188**
B.1 Hankel matrix 188
B.2 Hermite normal form 188
B.3 Howell normal form 189

C Gröbner bases applications **190**
C.1 Applications in algebra 190
C.2 Connections to geometry 192
C.2.1 A Word on Dimension 192

List of Figures

1.1 Intersections of low mutiplicity 2

1.2 Intersection of a parametric elliptic curve (curps) and quadrifolium when $t = 2$ 13

2.1 Monomials staircase of $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$ 47

2.2 Example of change to generic coordinates. 53

2.3 Shape Lemma for Example 1.1.1. 53

3.1 Monomials through untangling when μ is small. 72

3.2 Monomials through recursive untanting. 74

4.1 An initial segment T (green) and the monomials $\boldsymbol{E} = (y^4, xy^3, x^2y, x^4)$ (purple), with $s = 3$ and $\delta = 9$. 87

4.2 The first translate $\mathsf{T}_{\leftarrow 1}$ of T from Figure 4.1. 88

4.3 The shell of T from Figure 4.1. 89

4.4 The sets S, T and $\mathsf{U} = \mathsf{S} + \mathsf{T}$, with $\ell = h = 2$ and $n = m = 3$. 92

4.5 Two possible pavings with $d = 4$. 99

4.6 An improved paving. 99

List of Tables

2.1 Vis-à-vis affine varieties and affine schemes. 23

2.2 Observations on ring extensions and affine varieties 25

Chapter 1

Introduction

In the beginning, we have a zero-dimensional bivariate ideal I over a ring R; we shall describe the schemes of its quotient ring, $\operatorname{Spec} R/I$, while we shall limit the time we need to do so.

Solving polynomial equations is a fundamental prong of numerous problems. Notably, such equations are widely used in pure and applied mathematics, as intersection theory echoes in number theory, algebraic geometry, etc. It further plays a major role in a large body of related fields *e.g.*, computational geometry and topology, computer graphics and motion planning [45, 45, 152, 10, 78, 148].

1.1 Context

Let $\mathbb{A}$ be an integral domain, let $\mathbb{A} \hookrightarrow \mathbb{K}$ where $\mathbb{K}$ is a field, and consider an ideal $J = \langle f_1, \ldots, f_t \rangle \subseteq \mathbb{K}[\mathbf{x}]$ for $f_i \in \mathbb{A}[\mathbf{x}]$, where $\mathbf{x} = (x_1, \ldots, x_n)$. We refer to the solutions of the polynomial system formed by the f_i's as the *affine variety* of the ideal over the algebraic closure of $\mathbb{K}$, denoted $\bar{\mathbb{K}}$:

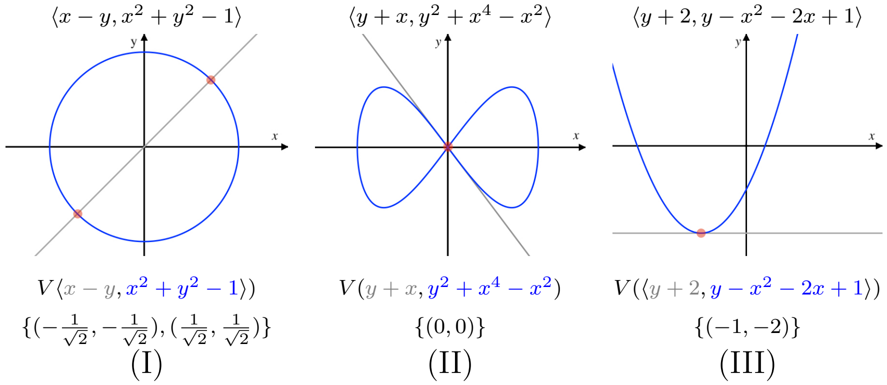
$$V(J) = \{\xi \in \bar{\mathbb{K}}^n, \quad f(\xi) = 0 \qquad \forall f \in J\}.$$

Natural incarnations of $\mathbb{K}$ in this document include but are not limited to: the rationals $\mathbb{Q}$, number fields, function fields, and their extensions. In the context of intersections of plane curves ($n = 2$), we denote $x_1 = x$ and $x_2 = y$.

Hereinafter, we focus on *zero-dimensional components of varieties* while drawing particular attention to the case with multiplicities (see Appendix C.2.1 for zero-dimensional ideals). Let I be a finite and non-empty intersection of primary components of J with $\mathbb{K}[x, y]/I$ of Krull dimension 0, then $V(I) \subseteq V(J)$ consists of a finite set of points (see Section 2.1.2 for primary components). From a geometric angle, when $n = 2$, these points correspond to intersections in the plane between curves $V(f_i)$,[1] where the f_i share no common factor.

Example 1.1.1: Intersections of multiplicity 1 and 2

Given the three below ideals in $\mathbb{R}[x, y]$ with two generators



$\langle x - y, x^2 + y^2 - 1\rangle$ $\langle y + x, y^2 + x^4 - x^2\rangle$ $\langle y + 2, y - x^2 - 2x + 1\rangle$

$V\langle x - y, x^2 + y^2 - 1\rangle)$ $\{(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\}$ (I)

$V(y + x, y^2 + x^4 - x^2)$ $\{(0, 0)\}$ (II)

$V(\langle y + 2, y - x^2 - 2x + 1\rangle)$ $\{(-1, -2)\}$ (III)

Figure 1.1: Intersections of low mutiplicity

the variety of each ideal corresponds to the intersection(s) (the red point(s) in the figure) of the two curves[a] described by each of the generators.

[a] Since we are observing the variety in the algebraic closure, these are complex curves in $\mathbb{A}^2(\mathbb{C})$, the complex affine plane, drawn in $\mathbb{R}^2$ for simplicity since no purely complex intersection exists.

Nonetheless, the point in the variety alone does not give the full portrait of an intersection; it conceals the *local structure* at the root (see Section 2.1.1 for local structure). In some cases, the curves may be singular at the intersection (Example 1.1.1, (II)), or the curves may be smooth but not transverse (Example 1.1.1, (III)).

[1] assuming $f_i \notin \mathbb{K}$

❧ **Definition 1.** *Let X be an algebraic variety, i.e. $X = V(I)$ with $I \subseteq \mathbb{K}[x_1, ..., x_n]$ an ideal, then X is **irreducible** if there do not exist two non-empty algebraic varieties Y and Z, such that $X = Y \cup Z$ but $X \neq Y$ and $X \neq Z$. A variety $X = V(I)$ is irreducible if and only if $\sqrt{I}$ is prime*[2]*. Otherwise, X is **reducible**. If $Z \subseteq X$ is an irreducible variety such that the only irreducible variety Y with $Z \subseteq Y \subseteq X$ is $Y = Z$ then Z is an **irreducible component***[3] *of X [45, §4.5].*

❧ **Definition 2.** *Let X be an irreducible affine variety in $\mathbb{A}^n(\mathbb{K})$, the n-dimensional affine space over an algebraically closed field $\mathbb{K}$. A point p in X is **smooth** if and only if the Zariski tangent space of X at p has dimension $\dim X$, equivalently the Jacobian $\left[\nabla f_1, \ldots, \nabla f_t\right]$ of*

$$\langle f_1, \ldots, f_t \rangle = I(X) := \{f \in \mathbb{K}[x_1, ..., x_n] | \forall p \in X, f(p) = 0\}$$

*evaluated at p has rank $n - dim(X)$ (see Appendix C.2.1 for dimension of a variety). Otherwise, the point is **singular** [85][§14].*

❧ **Definition 3.** *Let $I = \langle f_1, \ldots, f_t \rangle \subseteq \mathbb{K}[x_1, ..., x_n]$ be an ideal. A point p in $V(I)$ is **smooth** if*

- *p is not in the intersection of two irreducible components of X and;*
- *if $p \in Y \subseteq X$ then p is smooth in Y*

*Otherwise, p is **singular**.*

Both circumstances, singular points on a curve and non-transverse intersections, give rise to a non-trivial *multiplicity*. The multiplicity $\delta \in \mathbb{Z}$ of a solution $\xi = (\xi_1, \xi_2)$ can be obtained by looking at the dimension as a $\bar{\mathbb{K}}$-vector space of the local algebra at ξ [44]:

$$\delta_\xi = \dim_{\bar{\mathbb{K}}} \left(\bar{\mathbb{K}}[x, y]_{\langle x-\xi_1, y-\xi_2 \rangle} / I \bar{\mathbb{K}}[x, y]_{\langle x-\xi_1, y-\xi_2 \rangle} \right)$$

where $\bar{\mathbb{K}}[x, y]_{\langle x-\xi_1, y-\xi_2 \rangle}$ is the ring of rational functions $f(x, y)/g(x, y) \in \bar{\mathbb{K}}(x, y)$ that are well defined at ξ (see Section 2.1.1 for localisation). With some manipulations, one may equivalently prove that $\delta_\xi = \dim_{\bar{\mathbb{K}}} \bar{\mathbb{K}}[\![x, y]\!] / I_\xi \bar{\mathbb{K}}[\![x, y]\!]$, where

$$\bar{\mathbb{K}}[\![x, y]\!] = \{ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j} x^i y^j \mid a_{i,j} \in \bar{\mathbb{K}} \}$$

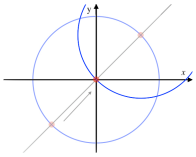
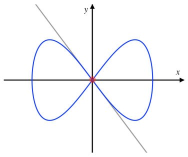
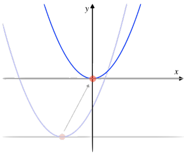
[2] An ideal $\mathfrak{p} \subseteq R$, where R is a commutative ring, is prime if and only if it is proper and $ab \in \mathfrak{p}$ implies a or b belongs to $\mathfrak{p}$.

[3] The irreducible components of $V(I)$ correspond to the primary components of I (see Section 2.1.2 for primary decomposition).

is the ring of power series and $I_\xi = \langle f(x+\xi_1, y+\xi_2) | f \in I\rangle$, which may simplify the evaluation of δ_ξ.

Example 1.1.2: Multiplicity

We may find the multiplicity of any root ξ from the previously defined varieties (Example 1.1.1) by translating the origin and looking at the basis of $\mathbb{C}[\![x,y]\!]/I_\xi\mathbb{C}[\![x,y]\!]$.

			
I_ξ	$\langle x-y,$ $(x-\frac{1}{\sqrt{2}})^2+(y-\frac{1}{\sqrt{2}})^2-1\rangle$	$\langle y+x,$ $y^2+x^4-x^2\rangle$	$\langle y,$ $y-x^2\rangle$
eliminating y using 1^{st} generator	$\langle x-y, x(2x-2\sqrt{2})\rangle$	$\langle y-x, x^4\rangle$	$\langle y, x^2\rangle$
Since $1/(2x-2\sqrt{2})$ exists in $\mathbb{C}[\![x,y]\!]$	$\langle x, y\rangle$		
basis of $\mathbb{C}[\![x,y]\!]/I_\xi\mathbb{C}[\![x,y]\!]$	$\{1\}$	$\{1,x,x^2,x^3\}$	$\{1,x\}$
	$\delta_{\{-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\}}=1$	$\delta_{\{0,0\}}=4$	$\delta_{\{-1,-2\}}=2$

For a zero-dimensional ideal $I \subseteq \mathbb{K}[x,y]$, $\sum_{\xi\in V(I)}\delta_\xi = \dim_{\bar{\mathbb{K}}}\big(\bar{\mathbb{K}}[x,y]/I\big)$ [44, §4.2, Corollary 2.5] and $\delta_\xi = 1$ if and only if p is smooth in $V(I)$ in which case we say the local structure is trivial.

While getting a set-theoretical description of $V(I)$, and possibly the multiplicity, is great, we want more. The *affine scheme* Spec R/I, defined in the next chapter, gives a more faithful description of the zero locus of I while preserving *the local structure*: a signature of the underlying primary ideal that characterises the regular functions defined at a point (see Section 2.1.1 for primary ideals and regular functions). This structure is notably useful when studying the local invariants of singular points [44, §4], algebraic operations on the roots [122], topology of curves or degree of polynomial maps [3, 63], analysis of ODEs and PDEs (bifurcation) [71] or local isomorphisms (see Section 2.1.1 for examples).

Our goal was to reduce the problem of describing $\operatorname{Spec} R/I$ and its localisation at an intersection while minimising the number of operations needed. To do so, we construct the Gröbner basis of I, defined in the next chapter, or the fibre of $I_{I(p)}$ under localisation, *i.e.* the $I(p)$-primary component of I, where $p \in V(I)$ and $I(p)$ is the maximal ideal that satisfies $V(I(p)) = p$.

Question 1.1.1. *Why do we get what we want?*

A Gröbner basis of I reduces the problem of finding the vanishing locus of I to a problem of factorisation (see Section 2.1.4 for examples). Crucially, the basis of the $I(p)$-primary component captures the local algebra's structure by accurately describing the localised quotient ring $R_{I(p)}/I_{I(p)}$.

1.2 The Results

Our goal is to describe the “algebraic” nature of intersections of plane curves. As mentioned, Gröbner bases of an ideal encode the position conjointly with the underlying structure of the local algebra at the intersections. Furthermore, the next chapter and Appendix C also discuss how they have proven useful for a large category of problems in algebra, algebraic geometry, topology, etc. Thence, we would like to address this question efficiently, which raises the question of whether the limit construction pattern, *e.g.* a $\mathfrak{m}$-adic expansion, is well tailored for the task. In particular, which convergence rate can be obtained? The question of whether quadratic convergence is achievable in the scheme of an $\mathfrak{m}$-adic limit construction for $\mathfrak{m} \subseteq \mathbb{A}$, a maximal ideal, is answered in this thesis.

We present an algorithm, GroebnerBasis, that features quadratic convergence to describe the local structure of the intersection of plane curves. Let $\mathbb{A}$ be a domain with $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal and $\mathbb{A} \hookrightarrow \mathbb{K}$ where $\mathbb{K}$ is a large field with a valuation (Archimedean or ultrametric), denoted $|\ |$ and defined in Section 5.3, and a fraction reconstruction algorithm with respect to the $\mathfrak{m}$-adic filtration (see Section 2.1.6 for fraction reconstruction). Further, let $\mathcal{F}$ be a finite subset of $\mathbb{A}[x, y]$. If $J = \langle \mathcal{F} \rangle \subseteq \mathbb{K}[x, y]$ is a zero-dimensional ideal and $\mathfrak{m} \in \mathbb{A}$ is a *good* maximal ideal of $\mathbb{A}$ with respect to $\mathcal{F}$ and $\mathcal{G}$, the lexicographic Gröbner basis of J, there exists an $\mathfrak{m}$-adic algorithm to construct the lexicographic Gröbner basis of J, or one of its primary components, with **quadratic convergence**.

The definition of a *good* maximal ideal for an $\mathfrak{m}$-adic expansion of a Gröbner basis is introduced in Chapter 4 and Chapter 5.

Let $I \subseteq J = \langle \mathcal{F} \rangle \subseteq \mathbb{K}[x, y]$ be ideals and I is zero-dimensional, the algorithm notably entails

- proving that the variety of the ideal generated the coefficients of $\mathcal{F}$ mod $\mathcal{G}$, where $\mathcal{G}$ is the parametric Gröbner basis of the stratum of $\mathbf{In}(I)$ (or Gröbner cell, defined in Section 2.1.7), has a smooth point that corresponds to the parameters of the Gröbner basis of I;
- giving an algorithm to find the canonical image of a Gröbner basis of interest in $\mathbb{A}/\mathfrak{m}[x, y]$ where $\mathfrak{m}$ is maximal in $\mathbb{A}$, the ambient ring of the original generators;
- bounding the coefficients of the reduced minimal basis of an ideal I based on an arbitrary generating set;
- characterising and proving the existence of a Zariski open for the set of maximal ideals $\mathfrak{m} \subseteq \mathbb{A}$ that are *good* to construct a $\mathfrak{m}$-adic limit.

To our knowledge, no bound were known to the coefficients of a reduced minimal Gröbner basis of an ideal. The last three points follow from a direct extension of results from Lazard [109] and Storjohann [149]: we prove that Hermite normal forms can be used to find a Gröbner basis, and the above results then follow from properties of these structured matrices.

The algorithm involves an original Newton iterator, LiftOneStep, for which we establish the arithmetic complexity over any field.

Theorem A (Chapter 4 - *cf.* Theorem 4.1.1 - **Bivariate Gröbner basis Newton iterator**). *Let $\mathbb{A}$ be a domain with $\mathbb{K}$ the fraction field of $\mathbb{A}$, $\mathcal{F} \subset \mathbb{A}[x, y]$ with $\deg f \leq d$ for all $f \in \mathcal{F}$ and let $J = \langle \mathcal{F} \rangle$ be an ideal in $\mathbb{K}[x, y]$. Let I be the intersection of some of the zero-dimensional primary components of J, with*

- $\delta = \dim_{\mathbb{K}} \mathbb{K}[x, y]/I$;
- *minimal, reduced lexicographic ($x \prec y$) Gröbner basis of I: $\mathcal{G}$;*
- $n_0 = \min\{i, y^i = in(f) \text{ for } f \in I\}$ *and* $m_s = \min\{i, x^i = in(f) \text{ for } f \in I\}$

.

If $\mathfrak{m} \subseteq \mathbb{A}$ is a good maximal ideal with respect to $\mathcal{G}$ and $\mathcal{F}$; there exists a Newton iterator to find $\mathcal{G} \mod \mathfrak{m}^{2^k}$, $k \in \mathbb{N}^+$, based in $\mathcal{G} \mod \mathfrak{m}$ using:

- *$O\tilde{}(s^2 n_0 m_s + |\mathcal{F}|\delta(d^2 + dm_s + s\delta + \delta^{\omega-1}))$ operations in $\mathbb{A}/\mathfrak{m}^{2^i}$, for $i = 1, \ldots, k$;*
- *$|\mathcal{F}|d^2 T_{2^i}$ steps for coefficient reduction, for $i = 1, \ldots, k$,*

where $O(n^\omega)$ is the asymptotic cost of matrix multiplication of a $n \times n$ matrix and T_{2^i} is the assumed time to reduce a coefficient of one f_j's modulo $\mathfrak{m}^{2^i}$ for $i \geq 0$.

Link to the past: *Until then, only linearly convergent algorithms were known [155, 135, 6] (see Section 2.2 for state of the art). Those iterative algorithms lift a Gröbner basis, its syzygies and its membership relations. In particular, they solve a multivariate system of unfixed size. Theorem A rely on modular arithmetic with a parametric basis with $O(\delta)$ parameters, due to an explicit bijection from Conca & Valla between a stratum and its associated moduli space, and solving a linear system in $\mathbb{K}$ ($\mathbb{A}/\mathfrak{m}^{2^i}$) of fixed size.*

For example, in the case of $\mathbb{A} = \mathbb{Z}$ this yields an optimal overall complexity vis-à-vis the size of the coefficients in the basis.

Theorem B (Chapter 5 - *cf.* Theorem 5.1.1 - **Complexity of building a lexicographic bivariate basis over $\mathbb{Q}$ and $k[\boldsymbol{t}]$).** *Let $\mathbb{A} = \mathbb{Z}$ (resp. $\mathbb{A} = k[\boldsymbol{t}]$), $\mathbb{K} = \bar{\mathbb{A}}$ and $\mathcal{F} \subset \mathbb{A}[x, y]$, then one can find the lexicographic ($x \prec y$) Gröbner basis $\mathcal{G}$ of a zero-dimensional ideal $I = \langle \mathcal{F} \rangle$, with high probability, in a number of binary operations (resp. operation in k) less than cubic in terms of $\dim_{\mathbb{K}} \mathbb{K}[x, y]/I$ and softly linear in the height (resp. degree in $\boldsymbol{t}$) of $\mathcal{G}$'s coefficients.*

Link to the past: *Buchberger's $\frac{3}{2}(|\mathcal{F}| + 2(d + 2)^2)^4$* ***operations in the field*** *[34] where $\delta \in O(d^2)$ by Bézout's theorem (see Section 2.2 for the state of the art).*

In our main results, the ring $\mathbb{A}$, and $\mathbb{K}$, have a natural notion of "size", *e.g.* the height when $\mathbb{A} = \mathbb{Z}$, *i.e.* $a \in \mathbb{Z}$ has height at most $h \in \mathbb{N}$ if $|a| \leq 2^h$, or the degree when $\mathbb{A} = k[t_1, \ldots, t_m]$. To cover the two types of valuations (*i.e.* Archimedean or ultrametric) with our examples, we also fill in the details for $k[\boldsymbol{t}]$.

The complexity obtained in Theorem A and Theorem B further entails:

- optimising modular arithmetic with a lexicographic Gröbner basis, which we address via paving of $\mathbb{N}^2$: the choice of generator in the basis to be used to reduce a monomial $x^a y^b$ where $(a, b) \in \mathbb{N}^2$;
- characterising a Zariski open of $\mathrm{GL}_2(\mathbb{K})$ (with action on $\mathbb{K}[x, y]$) that changes coordinates in a way that ensures the initial term ideal of a zero-dimensional ideal becomes Borel-fixed when $|\mathbb{K}|$ is sufficiently large (see Section 2.1.3 for initial term ideal).

The aforementioned action is the natural action for $A \in \mathrm{GL}_2(\mathbb{K})$ given by matrix-vector multiplication $A \begin{bmatrix} x \\ y \end{bmatrix}$ applied on each monomial of $f \in \mathbb{K}[x, y]$. An ideal I is said Borel-fixed if I is invariant under the Borel group, *i.e.*

$$I = \left\{ Af \mid f \in I, A = \begin{bmatrix} *, 0 \\ *, * \end{bmatrix} \in \mathrm{GL}_2(\mathbb{K}) \right\}.$$

This definition of Borel differs from the conventions, *e.g.* [61, Chapter 15], which uses upper-triangular matrices; this is due to our choice of monomial ordering ($x \prec y$). Based on a result from Pardue [134] [61, 15.9], this allows us to evaluate the probability of being able to reduce the complexity in Theorem A (notably the terms with n_0 and m_s).

The method of Theorem B can also be extended to a primary component of the ideal. In particular, we present the complexity analysis for the $\langle x, y \rangle$-primary component.

Theorem C (Chapter 5 - *cf.* Theorem 5.1.2 - **Complexity of finding the $\langle x, y \rangle$-primary component**). *Let $\mathbb{A} = \mathbb{Z}$ (or $\mathbb{A} = k[\boldsymbol{t}]$), $\mathbb{K} = \bar{\mathbb{A}}$ and $\mathcal{F} \subset \mathbb{A}[x, y]$, then one can find the lexicographic ($x \prec y$) Gröbner basis $\mathcal{G}$ of I the $\langle x, y \rangle$-primary component of $\langle \mathcal{F} \rangle$, with high probability, in a number of binary operations (operation in k) that is less than cubic in terms of* $\dim_{\mathbb{K}} \mathbb{K}[x, y]/I$ *and softly linear in term of the height (degree in $\boldsymbol{t}$) of $\mathcal{G}$'s coefficients.*

Theorem B and Theorem C are probabilist in the sense that they depend on the choice of a good maximal ideal and a generic change of coordinates.

In order to use the result of Theorem C for intersections that are not at the origin, we directly extend an isomorphism of change of basis from van der Hoeven *&* Lecerf [91] to the bivariate case (Chapter 3).

Theorem D (Chapter 3 - *cf.* Proposition 3.4.1 - Bivariate untangling)**.** *Assume* $\mathfrak{p}$ *is a maximal ideal in* $\mathbb{K}[x,y]$ *and* I *is a* $\mathfrak{p}$*-primary zero-dimensional ideal in* $\mathbb{K}[x,y]$*, with* $\mathbb{K}$ *perfect of characteristic at least* $\deg(I)$*.*

Let $\tilde{\mathfrak{p}}$ *be the image of* $\mathfrak{p}$ *through the isomorphism* $\mathbb{K}[x,y] \cong \mathbb{K}[x',y']$*, let* α_1, α_2 *be the residue classes of* x', y' *in* $\mathbb{F} := \mathbb{K}[x',y']/\tilde{\mathfrak{p}}$ *and let* J *be the primary component of* $I \cdot \mathbb{F}[\xi_1, \xi_2]$ *at* (α_1, α_2)*. Finally, let* J' *be the image of* J *through* $(\xi_1, \xi_2) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$*. Then, there exists an* $\mathbb{K}$*-algebra isomorphism*

$$\pi_{\mathfrak{p},J'} : \mathbb{K}[x,y]/I \to \mathbb{F}[\xi_1, \xi_2]/J' \tag{1.2.0.1}$$

given by $(x,y) \mapsto (\xi_1 + \alpha_1, \xi_2 + \alpha_2)$ *where* J' *is* $\langle \xi_1, \xi_2 \rangle$*-primary.*

By this change of bases, we obtain a primary component at the origin isomorphic to the original $\mathfrak{p}$-primary component that simplifies several operations, particularly modular arithmetic with a basis. In counterpart, we have to work on an ambient field larger than the one of departure, *i.e.*, we adjoin algebraic values.

Example 1.2.1: Trivial case of tangling

Let $I = \langle x - y, x^2 + y^2 - 1 \rangle \cap \langle y + x, y^2 + x^4 - x^2 \rangle \subseteq \mathbb{Q}[x,y]$, from Example 1.1.1 (I) and (II), and let $\mathfrak{p} = \langle x - y, x^2 + y^2 - 1 \rangle$. To focus on the local structure of the points in $V(\mathfrak{p})$, we could localise $\mathbb{Q}[x,y]_{\mathfrak{p}}/I_{\mathfrak{p}}$: this describes the two points tangled with their local structure (trivial).

To move the points in $V(\mathfrak{p})$ at the origin, one could define α_1, α_2 to be the residue classes of x', y' in $\mathbb{F} := \mathbb{Q}[x',y']/\tilde{\mathfrak{p}}$ where $\tilde{\mathfrak{p}}$ be the image of $\mathfrak{p}$ through the isomorphism $\mathbb{K}[x,y] \cong \mathbb{K}[x',y']$. Now the $\langle \xi_1 - \alpha_1, \xi_1 - \alpha_2 \rangle$ primary component of $I \cdot \mathbb{F}[\xi_1, \xi_2]$ can be translate at the origin that we denote J'.

Here $J' = \langle \xi_1, \xi_2 \rangle \subseteq \mathbb{Q}(\sqrt{2})[x,y]$ and $\dim_{\mathbb{Q}(\sqrt{2})} \mathbb{Q}(\sqrt{2})[\xi_1, \xi_2]/J' = 1$, *i.e.* the multiplicity at the origin is 1, which is expected given the isomorphism $\mathbb{Q}(\sqrt{2})[\xi_1, \xi_2]/J' \cong \mathbb{Q}[x,y]/\mathfrak{p} \cong \mathbb{Q}[x,y]_{\mathfrak{p}}/I_{\mathfrak{p}}$ (see Example 1.1.2 for the multiplicity of the points in $\mathfrak{p}$).

❧ Example 1.2.2: Another trivial case of tangling and ring extension

Let $I = \langle y + 2, y - x^2 - 2x + 1\rangle \subseteq \mathbb{Q}[x, y]$ from Example 1.1.1 (III). To use Theorem C on the points that are not at the origin, one could define $\mathfrak{p} = \langle x-1, y-2\rangle$ and let α_1, α_2 be the residue classes of x', y' in $\mathbb{F} := \mathbb{Q}[x', y']/\tilde{\mathfrak{p}}$ where $\tilde{\mathfrak{p}}$ be the image of $\mathfrak{p}$ through the isomorphism $\mathbb{K}[x, y] \cong \mathbb{K}[x', y']$. Now the $\langle \xi_1 - \alpha_1, \xi_1 - \alpha_2\rangle$ primary component of $I \cdot \mathbb{F}[\xi_1, \xi_2]$ can be translate at the origin that we denote $J' = \langle \xi_1^2, \xi_2\rangle \subseteq \mathbb{Q}[x, y]/I$. We observe we do get $\mathbb{Q}[\xi_1, \xi_2]/\langle \xi_1^2, \xi_2\rangle \cong \mathbb{Q}[x, y]/\langle y + 2, y - x^2 - 2x + 1\rangle$ and in particular, the local structure of the point is preserved.

Remark 1.2.1. The above example allows us to visualise that more than one point can be moved at the origin and the local structures are preserved. In Example 1.2.1, the method of Theorem C would be superfluous, as the local structure of the points is trivial. Chapter 3 present examples that are still simple but more interesting from an algorithmic and local structure perspective.

We complete the complexity analysis for the maps in Theorem D under the hypothesis that the resulting ideal is monomial, thereby we partially address

- efficiently change basis to move a primary component to the origin;

but the general case is left as future work.

1.2.1 Examples

❧ Example 1.2.3: Small example

Let

$$\begin{aligned} f_1 &= (2y) + (3x) \\ f_2 &= (2y)^2 + (3x)^4 - (3x)^2 \\ g_1 &= x + 2 \\ g_2 &= x - y^2 - 2y + 1 \end{aligned}$$

and $\mathcal{F} = \{f_1 g_1, f_1 g_2, f_2 g_1, f_2 g_2\}$. The ideal $I = \langle \mathcal{F} \rangle \subseteq \mathbb{C}[x, y]$ is the product of

- $I_1 = \langle f_1, f_2 \rangle$ (Example 1.1.1 II stretched, with unique solution $(0, 0)$ of multiplicity 4) and
- $I_2 = \langle g_1, g_2 \rangle$ (Example 1.1.1 III after a change of coordinates

  $$\begin{aligned} x &\mapsto y \\ y &\mapsto x, \end{aligned}$$

  with unique solution $(-2, -1)$ of multiplicity 2).

Since $V(I_1) \cap V(I_2) = \emptyset$, the two ideals are coprime and $V(I) = V(I_1) \cup V(I_2)$, and for $p \in V(I_i)$, p inherits its local structure from I_i.

First, we want to find an *approximation* of $\mathcal{G}$, the lexicographic $(x \prec y)$ Gröbner basis of I, *e.g.* $\varphi_{p,1}(g)$ for $g \in \mathcal{G}$ where $\varphi_{p,i} : \mathbb{Z}_{\langle p \rangle}[x, y] \to \mathbb{Z}/p^i\mathbb{Z}[x, y]$ is the canonical projection. Here, the ideal $5\mathbb{Z}$ is a *good* maximal ideal with respect to $\mathcal{F}$ and $\mathcal{G}$ (see Chapter 5) in particular $\mathcal{G} \subset \mathbb{Z}_{\langle 5 \rangle}[x, y] \subset \mathbb{Q}[x, y]$ so $\varphi_{5,*}$ is well defined on elements of $\mathcal{G}$. Let π_d denote the $\mathbb{K}[x]$-module isomorphism

$$\begin{array}{rccc} \pi_d : & \{f \in \mathbb{Z}/5\mathbb{Z}[x, y] \mid \deg_y(f) < d\} & \to & \mathbb{Z}/5\mathbb{Z}[x]^d \\ & \sum_{i=0}^{d-1} a_i y^i & \mapsto & [a_{d-1}, \cdots, a_0]^\top. \end{array}$$

Similarly to what Lazard did for 2 generators [109], we define $S = [S_{\mathcal{F}(1)}, \ldots, S_{\mathcal{F}(4)}]$ where $S_f = [\pi_8(\varphi_{5,1}(y^4 f)), \ldots, \pi_8(\varphi_{5,1}(yf)), \pi_8(\varphi_{5,1}(f))]$ and we find its Hermite normal form :

$$\boldsymbol{H} = \begin{bmatrix} \pi_8(y^7 + 3y + 3x^4 + 2x)^\top, \\ \pi_8(y^6 + y + 3x^4 + 4x)^\top, \\ \pi_8(y^5 + x^4)\pi_8(y^4 + 4y + x)^\top, \\ \pi_8(y^3 + 2y + x^4 + 4x^3 + 3x)^\top, \\ \pi_8(y^2 + 2y + x^4 + 4x^2 + 3x)^\top, \\ \pi_8((x+2)y + 4x^2 + 3x)^\top, \\ \pi_8(x^5 + 2x^4)^\top \end{bmatrix}^\top .$$

(see Appendix B.2 for Hermite normal form).

The preimage of the columns of $\boldsymbol{H}$ form a lexicographic ($y \prec x$) Gröbner basis of an ideal of $\mathbb{Z}/5\mathbb{Z}[x, y]$ for which

$$\{\varphi_{5,1}(g) \mid g \in \mathcal{G}\} = \{y^2 + 2y + x^4 + 4x^2 + 3x, (x+2)y + 4x^2 + 3x, x^5 + 2x^4\}$$

is the reduced minimal form. Applying LiftOneStep once we obtain $\varphi_{5,2}(\mathcal{G})$

$$\left| \begin{array}{l} y^2 + 2y + x^4 + 4x^2 + 3x \\ yx + 2y + (4 + 5^2)x^2 + 3x \\ x^5 + 2x^4 \end{array} \right.$$

After 4 repetitions and rational reconstruction we start stabilizing with at $\mathcal{G}$, *i.e.*,

$$\left| \begin{array}{l} y^2 + 2y + x^4 - \frac{9}{4}x^2 + 3x, \\ xy + 2y + \frac{3}{2}x^2 + 3x, \\ x^5 + 2x^4 \end{array} \right.$$

❧ Example 1.2.4: Intersection of an elliptic curve and a quadrifolium

Let $R = \mathbb{C}[t]$ and let

$$f_1 = (t^2-1)y^3 + (t^3-1)x^2$$
$$f_2 = (x^2+y^2)^3 - 4(t^2-t+1)^2x^2y^2$$

and let $\mathcal{F} = \{f_1, f_2\}$ and $J = \langle \mathcal{F} \rangle \subseteq \mathbb{C}(t)[x,y]$ be an ideal. Here $(0,0) \in V(J)$ is singular; let I be the $\langle x, y\rangle$-primary component of J. For example, when $t \notin \{0, \pm 1, \zeta_3, \zeta_3^2\}$, for ζ_3 the third root of unity, $V(J)$ is the intersection of an elliptic curve (grey) and a quadrifolium (blue)

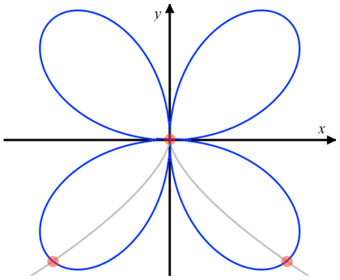


Figure 1.2: Intersection of a parametric elliptic curve (curps) and quadrifolium when $t = 2$

The ideal $\langle t \rangle$ is a *good* maximal ideal (definition in Chapter 5) with respect to $\mathcal{F}$ and the lexicographic $(x \prec y)$ Gröbner basis of I. Define $L = (R/\langle t\rangle)[x]/\langle x^{18}\rangle \cong \mathbb{C}[x]/\langle x^{18}\rangle$ and

$$\pi_n : \{f \in L[y] \mid \deg_y(f) < n\} \to L^n$$

the $\mathbb{K}[x]$-module isomorphism which maps $a_0 + \cdots + a_{n-1}y^{n-1}$ to the vector $[a_{n-1}, \cdots, a_0]^\top$. The Howell normal form (see Appendix B.3 for definition) of

$$S = [\pi_{12}(y^6\bar{f}_1), \ldots, \pi_{12}(y\bar{f}_1), \pi_{12}(\bar{f}_1), \pi_{12}(y^6\bar{f}_2), \ldots, \pi_{12}(y\bar{f}_2), \pi_{12}(\bar{f}_2)],$$

where $\overline{f_i}$ is the canonical projection of f_i in $L[y]$, is

$$\boldsymbol{H} = \left[\pi_{12}(y^{11}), \ldots, \pi_{12}(y^5), \pi_{12}(y^4 + x^2y), \pi_{12}(y^3 + x^2), \pi_{12}(x^2y^2), \pi_{12}(x^4y), \pi_{12}(x^4)\right].$$

The π_{12}-fiber of the columns of $\boldsymbol{H}$ forms a lexicographic ($x \prec y$) Gröbner basis of an ideal of $R/\langle t \rangle[x,y]$ for which $\mathcal{G} = \{y^3 + x^2, y^2x^2, x^4\}$ is a reduced minimal form.

After 4 repetitions of LiftOneStep and a fraction reconstruction, we already obtain the Gröbner basis of I, *i.e.*,

$$\left|\begin{array}{l} y^3 + \frac{t^2+t+1}{t+1}x^2, \\ y^2x^2, \\ x^4. \end{array}\right.$$

1.2.1.0.1 Leitfaden

Shall we ,

This document is structured as follows: Chapter 2 features the *mise en place*; we review the motivations and the theoretical concepts pilar to the outcomes. We recall that finding a Gröbner basis of a $I(p)$-primary components, for a point p, is equivalent to describing the quotient ring R/I localised at p where connections to algebraic geometry can be seen through scheme theory. We also review the definition and motivations of Gröbner bases. The chapter ends by showing a brief overview of the literature; each subsequent chapter is accompanied by its own specific literature review.

Chapter 3 proves Theorem D to move primary components to the origin while preserving the local structure. The isomorphism, inspired by van der Hoeven *&* Lecerf [91], called tangling and untangling, separates (untangling) or combines (tangling) the local structure from the points. The operator which isolates the local structure moves the points at the origin. This change of coordinates aims to simplify the arithmetic of modular operations with a Gröbner basis that, in due course, we use repeatedly. We review the univariate case to offer an algorithm with slightly better complexity (see Proposition 3.3.1). Then we present a divide-and-conquer algorithm to handle bivariate monomials ideals (see Propositions 3.4.2 and 3.4.3). After the change of basis, the resulting ideal is defined over an extension of the original ambient ring. Hence, to serve more than a theoretical purpose, we keep general rings and extensions throughout the main line of the following chapters. Tangentially, we further prove an improvement to the complexity bound for finding elements in non-square-free linear recurrence sequences using Fiduccia's algorithm [70].

Chapter 4 proves Theorem A which makes possible an $\mathfrak{m}$-adic construction that features a quadratic convergence to construct a lexicographic Gröbner basis of an ideal $I \subseteq \mathbb{K}[x,y]$ with a generating set in $\mathbb{A}[x,y] \subset \mathbb{K}[x,y]$, where $\mathfrak{m} \subseteq \mathbb{A}$ is a maximal ideal. The algorithm relies on an explicit bijection ϕ between a stratum (Section 2.1.7) and the corresponding moduli space from [40] to exploit the simple property that for two ideals I, J over a given ring then $(I + J)/J = 0$ (*i.e.* $a \mod J \equiv 0$ for all $a \in I$) if and only if $I \subseteq J$. The key idea is to define a parametric basis $\mathcal{G}$ for a stratum, a set of ideals that contains I. The intuition is that when replacing the parameters in $\mathcal{G}$ with $\phi(I)$, we can find a Gröbner basis of I. When projecting a generating set of J in $\mathbb{K}[x, y]/\mathcal{G}$, the inclusion of the ideals statement tells us that $\phi(I)$ is a zero of the ideal of the coefficients. In Chapter 4, we prove that the above ideal of coefficients vanishes at $\phi(I)$ with multiplicity 1. As a result, a Newton iteration can be applied to construct the image in the $\mathfrak{m}$-adic completion. As the key idea is to perform modular arithmetic with a generic Gröbner basis, we address the complexity via a segmentation of $\mathbb{N}^2$, which we call a *paving*, to base our choice of generators to reduce each monomial.

In Chapter 5, we present the main algorithm. We also prove some general properties of a basis of an ideal. When there exists a valuation (Archimedean or ultrametric) for a large base field, we prove the existence of *good maximal ideals* and characterise the growth of the coefficients (*i.e.*: bound the valuation) in the representation from which follows the number of iterations required to recover the basis. Both proofs follow the results from [149] and the property that, in generic coordinates, a Gröbner basis can be read out of the Hermite normal form of a matrix created by the generators (Proposition 5.2.1 in Chapter 5). Inspired by Lazard [109], this usage of the Hermite normal form also originates an approximation in $\mathbb{A}/\mathfrak{m}[x, y]$ for our basis to start a limit construction. Over a field of large characteristics, we also prove that the initial term ideals of zero-dimensional ideals in generic coordinates are Borel-fixed when $|\mathbb{K}|$ is sufficiently large. The conclusion could be deduced from [144], but not in a effective manner, in the sense that it does not give description of the Zariski open. Galligo [72], Bayer-Stillman [11] and Pardue [134] (summarised in [61]) offered a description of generic monomial ideals and proved that they are Borel-fixed for homogeneous ideals. We adapt the key ideas to show the result under a different hypothesis, *i.e.* $|V(I)| < \infty$ instead of I being homogeneous. In particular, this description leads to a nice property of a $\mathbb{K}$-basis of the quotient ring $(\mathbb{K}[x_1, \ldots, x_n]/I)$, which may be used in complexity analysis; based on the result from [134], there exists a *well distributed* monomial basis for of R/I for I. Concretely, under the assumption of genericity for the coordinates, it implies for all but one variable x_i: a better bound than the degree of the I for $\min\{a_i : x_i^{a_i} \in \mathbf{In}(I)\}$ where I is a zero-dimensional ideal. This chapter concludes by exposing the culminating complexity analysis of Theorem B and Theorem C.

Chapter 6 presents some open questions we wish to address. We also discuss different implications for some choices of I in Theorem A oriented by the Chinese remainder theorem but viewed under a complexity and algebraic geometry perspective.

The appendices summarise some definitions of general algebra concepts, namely, the exterior product (Appendix A), used in Chapter 5 to prove the Borel-fixed property, and special matrices (Appendix B), used in Chapter 3 and Chapter 5 to prove the condition on the maximal ideals, a bound on the growth of the coefficients and to find the canonical image of the basis in $\mathbb{A}/\mathfrak{m}$ for $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal.

Chapter 2

Observations

The problem arises from classical algebraic geometry and intersection theory. Furthermore, four theoretical concepts are key to this thesis: primary decompositions, Gröbner bases, limits and moduli space of strata.

2.1 Review of the key concepts

2.1.1 Algebraic varieties and Affine Schemes

To whom shares my interest in schemes,

❧ Preamble: We now present the geometric context in which this project occurs. The idea is to illustrate the contrast between algebraic varieties and affine schemes. This should highlight the advantages of using methods such as Gröbner bases to describe the quotient ring of an ideal by emphasising how the quotient ring gives more than the varieties. A thorough understanding of these concepts is not required to benefit from the rest of this document, although it offers great insight. ☙

2.1.1.0.1 Algebraic Varieties Let $R = \mathbb{K}[x_1, \ldots, x_n]$ be a commutative polynomial ring over an arbitrary field $\mathbb{K}$. Given an ideal $I = \langle f_1, \ldots, f_s \rangle \subseteq R$, the affine variety of I, denoted $V(I)$, is the vanishing locus of I in $\bar{\mathbb{K}}^n$, that is, the set of the common roots of the f_i's, which we can make a topological space.

❧ **Definition 4.** *Let X be a set, a* ***topology*** *on X is a collection $\mathcal{T}$ of subsets of X, which we call open sets, such that*

- *X and $\emptyset$ are in $\mathcal{T}$;*
- *$\mathcal{T}$ is closed under finite intersection;*
- *$\mathcal{T}$ is closed under union [29, §1.2].*

❧ **Definition 5.** *A* ***topological space*** *$(X, \mathcal{T})$ is a set X, possibly empty, with a topology $\mathcal{T}$ on X.*

A subset $C \subseteq X$ is **closed** in a topological space $(X, \mathcal{T})$ if there exists $U \in \mathcal{T}$ such that $C = X \setminus U$.

❧ Example 2.1.1: Affine space

The affine space in n dimensions over a field $\mathbb{K}$, $\mathbb{A}^n(\mathbb{K})$ or $\mathbb{K}^n$, is a topological space where the closed sets are the $V(I)$'s for all ideals the $I \subseteq \mathbb{K}[x_1, \dots, x_n]$.

❧ Example 2.1.2: Projective space

The projective space in n dimension over a field $\mathbb{K}$, $\mathbb{P}^n(\mathbb{K})$ or $\mathbb{P}^n$ defined as

$$\mathbb{P}^n := \{(p_1 : \dots : p_n) \mid p_i \in \mathbb{K} \text{ and } (p_1 : \dots : p_n) \neq (0 : \dots : 0)\}$$

with $(p_1 : \dots : p_n) = (kp_1 : \dots : kp_n)$ for all $k \in \mathbb{K}^\times$, is a topological space where the closed sets are the $V(I)$'s for all homogeneous ideals[a] $I \subseteq \mathbb{K}[x_1, \dots, x_n]$.

[a] An ideal I is homogeneous whenever $f \in I$ ensures that the homogeneous components of f, *i.e.*, the sum of the monomials of f of a given degree, are also in I. This happens if and only if there exists a generating set $\{f_1, \dots, f_t\}$ of I such that for all i: all monomials in f_i have degree a_i for some $a_i \in \mathbb{N}$

❧ Example 2.1.3: Variety of an ideal

For $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ an ideal, $V(I) \subseteq \mathbb{A}^n(\mathbb{K})$ is a topological space, where the closed sets are the $V(J)$'s for all ideals $I \subseteq J$.

Over an algebraically closed field, Hilbert proved a useful theorem relating algebraic varieties and ideals, the Nullstellenstaz, which is used sporadically in the document.

Theorem 2.1.1. *(The Nullstellensatz) [45, §5] Let $\mathbb{K}$ be an algebraically closed field, then we have a correspondence*

$$\{\textit{affine subvarieties of } \mathbb{A}^n(\mathbb{K})\} \leftrightarrow \{I \subseteq \mathbb{K}[x_1, \dots, x_n] | I = \sqrt{I}\}$$

with in particular $V(I) = \emptyset$ if and only if $I = \langle 1 \rangle$ and $V(J)$ is a point in $\mathbb{A}^n(\mathbb{K})$ if and only if J a maximal ideal.

The radical ideal that corresponds to a variety $X \subseteq \mathbb{A}^n(\mathbb{K})$ for a field $\mathbb{K}$ is

$$I(X) := \{f \in \mathbb{K}[x_1, ..., x_n] \mid \forall p \in X, f(p) = 0\}.$$

There is also a correspondence with nilpotent-free rings over an algebraically closed field $\mathbb{K}$ via the coordinate ring $\mathbb{K}[x_1, \dots, x_n]/I$ for I a radical ideal [62].

2.1.1.0.2 Localisation In some situations, it can be useful to focus on a subset of a variety. This can be done via localisation.

❧ **Definition 6.** *Let $\mathbb{A}$ be a ring, D a multiplicatively closed subset of $\mathbb{A}$ and $\mathcal{M}$ an $\mathbb{A}$-module. Then the **localisation** of $\mathcal{M}$ away from D is defined as $D^{-1}\mathcal{M} := \{m/d \mid m \in \mathcal{M} \text{ and } d \in D\}/\sim$, where given $m_1, m_2 \in \mathcal{M}$ and $d_1, d_2 \in D$ we say $m_1/d_1 \sim m_1/d_2$ if there exists $d \in D$ such that $d(m_1/d_1 - m_2/d_2) = 0$ [7][§3].*

If $\mathcal{M}$ is generated by $\mathcal{G}$ then $D^{-1}\mathcal{M}$ is a $D^{-1}\mathbb{A}$ module generated by $\mathcal{G}$. When localising an ideal $I \subseteq R$ at a prime ideal J with $X = V(J)$ (a point, a line, a hypersurface...), we equivalently write I_X or $I_J := [R \setminus J]^{-1}I$. When studying the geometries of objects, localisation is aptly named by focusing exclusively on what happens in the vicinity of X; everything away from X is perceived as nonexistent – out of sight, out of localisation.

2.1.1.0.3 Affine schemes As a generalisation of the concept of algebraic varieties, Alexander Grothendieck introduced the notion of schemes which comes with a broader correspondence [62, §I.1]

$$\{\text{affine scheme}\} \leftrightarrow \{\text{commutative rings with identity}\}.$$

We will focus on the rings of the form $\mathbb{K}[x_1, \dots, x_n]/I$ for all ideals $I \subseteq \mathbb{K}[x_1, \dots, x_n]$. The notion of scheme is defined using the spectrum of a ring and sheaves.

❧ **Definition 7.** *The* ***spectrum*** *of a ring R,* $\operatorname{Spec} R$*, is the set of prime ideals in* R[1]*.*

❧ **Definition 8.** *A* ***sheaf*** $\mathscr{F}$ *of a topological space X is a family of sets on all the open sets $U \subseteq X$, noted $\mathscr{F}(U)$ and called sections, that not only satisfies the following conditions on the restrictions (presheaf):*

- *If $V \subseteq X$ is open with $V \subseteq U$, and if $\mathscr{F}(U)$ is a section on U, then there exists a restriction of $\mathscr{F}(U)$ to $\mathscr{F}(V)$ where $\mathscr{F}(U) \mid_V$ is the image and where $\mathscr{F}(U) \mid_U = \mathscr{F}(U)$;*

- *If $V, W \subseteq X$ are open with $W \subseteq V \subseteq U$, then $(\mathscr{F}(U) \mid_V) \mid_W = \mathscr{F}(U) \mid_W$*

but also glues well on the intersection of open sets: whenever $X = \bigcup U_i$ where $U_i \subseteq X$ are open (open cover) and given $f_i \in \mathscr{F}(X) \mid_{U_i}$ for all i such that $f_i \mid_{U_i \cap U_j} = f_j \mid_{U_i \cap U_j}$ then there exists $f \in \mathscr{F}(X)$ such that $f \mid_{U_i} = f_i$ for all i. It further requires uniqueness on the open covers: if $\mathscr{F}$ and $\mathscr{G}$ are sheaves on X and $\mathscr{F}(U_i) = \mathscr{G}(U_i)$ for all U_i is a open cover then $\mathscr{F} = \mathscr{G}$ [154, §2.2].

A sheaf $\mathscr{F}$ is a sheaf of rings if the family on the open set are rings and the restrictions are ring morphisms.

❧ **Definition 9.** *Let $\mathscr{F}_1$ and $\mathscr{F}_2$ be sheaf on X then a map on the family of sets $\rho_U : \mathscr{F}_1(U) \to \mathscr{F}_2(U)$, for all the open set $U \subseteq X$, is a* ***morphism of sheaves*** *if*

$$\rho_V(\mathscr{F}_1(U) \mid_V) \to \rho_V(\mathscr{F}_2(U) \mid_V)$$

for all pair of open sets $V \subseteq U$ [62, I.1]. If $\mathscr{F}$ is a sheaf of rings, then the morphism $\mathscr{F}_1(U) \to \mathscr{F}_2(U)$ is a ring morphism.

[1] An ideal $\mathfrak{p} \subseteq R$, where R is a commutative ring, is prime if and only if it is proper and $ab \in \mathfrak{p}$ implies a or b belongs to $\mathfrak{p}$.

❧ **Definition 10.** *An **affine scheme** is a topological locally ringed space $(X, \mathscr{O}_X)$, which means it consists of*

(a topological space) *:* *$X = \operatorname{Spec} R$, i.e. the points are primes in some ring R, with the topology where the closed sets are*

$$\mathscr{V}(S) = \{[I] \in \operatorname{Spec} R | S \subset I\} \subseteq \operatorname{Spec} R$$

for all subsets $S \subseteq R$ [2]*– this topology is called the **Zariski** topology. For any open $U \subseteq X$, $(U, \mathcal{O}_X(U))$ is also a scheme called the open subscheme of X [62, §I.1.2] ;*

and

(a sheaf of rings): *the sheaf of rings of regular functions $\mathscr{O}_X$ on X, called structure sheaf. The structure sheaf can be defined under localisation for $U \subseteq X$ an open set. This is denoted by $\mathscr{O}_X(U)$ and defined as the regular functions on U that can be inverted in the neighbourhood U. The regular functions on U are fractions f/g where*

- *$f, g \in R$*
- *$f(p), g(p)$ for $p \in U$ is the equivalence class of f and g in R/p*
- *g is nowhere vanishing on U.*

[86, §2] [62, §I.1].

Remark 2.1.1. Let $\mathbb{A}$ be a ring, since the only elements of $\mathbb{A}$ that do not vanish at any point of $X = \operatorname{Spec} \mathbb{A}$ are also invertible in $\mathbb{A}$, regular functions on X are equivalently written as elements of $\mathbb{A}$ [62, §I.1.1].

To denote an affine scheme, more often than not, we simply write X or $\operatorname{Spec} \mathbb{A}$, when $X = \operatorname{Spec} \mathbb{A}$ for a known ring $\mathbb{A}$, instead of $(X, \mathscr{O}_X)$. In the case of an affine scheme $X = \operatorname{Spec} R/I$, the global sections of the structure sheaf, *i.e.* $\mathscr{O}_X(X)$, is

[2] Note the properties: $\mathscr{V}(S) = \mathscr{V}((S))$, *i.e.* the ideal generated by S, for ideal I, J; $\mathscr{V}(IJ) = \mathscr{V}(I) \cup \mathscr{V}(J)$, $\mathscr{V}(I + J) = \mathscr{V}(I) \cap \mathscr{V}(J)$ and $\mathscr{V}(I) \subseteq \mathscr{V}(J)$ if and only if $\sqrt{I} \supseteq \sqrt{J}$ [86, §II.2, Lemma 2.1].

isomorphic to R/I and the closed sets are the $\mathscr{V}(JR/I)$ where $J \subseteq R$ is an ideal containing I. If X is of finite type over $\mathbb{K}$, that is if $X = \operatorname{Spec} \mathbb{K}[x_1, \dots, x_n]/I$ with $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ an ideal, the regular functions on the open sets $U \subseteq X$ can be written as f/g where f, g are polynomials in n variables that maps points in U to element in $\mathbb{K}$, where g does not vanish on U, and the $\mathscr{V}(S)$ for $S \subset R$ are the points where S *vanishes*.

❧ Example 2.1.4: The affine line [86, §I.1, Example 1.1.1]

Let $\mathbb{K} = \bar{\mathbb{K}}$ and $X = \operatorname{Spec} \mathbb{K}[x]$: the affine line. The points (primes) of X are $\langle 0 \rangle$ and the maximal ideals $\langle x - a \rangle$ for $a \in \mathbb{K}$ and the closed sets of the X are

- the *vanishing* of any polynomial $f \in \mathbb{K}[x]$, $f = \prod_{p \in P}(x - p)$ with $P \subset \mathbb{K}$ a finite set, $\mathscr{V}(\langle f \rangle) = \{[I] \in \operatorname{Spec} \mathbb{K}[x] \mid [f]_I = [0]_I\} = \{\langle x - p \rangle \mid p \in P\}$;
- and, naturally, the empty set and the full set.

Here $\mathscr{O}_X(X) = \mathbb{K}[x]$ and for $f \in \mathbb{K}[x]$ with $U = X \setminus \mathscr{V}(\langle f \rangle)$ then $\mathscr{O}_X(U) = \{\frac{g}{f^i} \mid i \in \mathbb{N}, g \in \mathbb{K}[x]\} = \mathbb{K}[x]_f$ [86, §II.2 Proposition 2.2 and Exercise 2.1].

❧ Example 2.1.5

Let $X = \operatorname{Spec} \mathbb{K}[x]/\langle x^i \rangle$ for $i \in \mathbb{N}^+$ then there only exists one point: $\langle x \rangle$. The closed sets are $\{\emptyset, \operatorname{Spec} \mathbb{K}[x]/\langle x^i \rangle\}$. Here $\mathscr{O}_X(X) = \mathbb{K}[x]/\langle x^i \rangle$.

❧ Example 2.1.6

Let $I = (\langle x - a, y - b \rangle \cap \langle x^i, y^j \rangle) \subseteq \mathbb{K}[x, y]$ where $a \neq 0$, $b \neq 0$ and $X = \operatorname{Spec} \mathbb{K}[x, y]/I$, then the points are $\{\langle x, y \rangle, \langle x - a, y - b \rangle\}$ and closed sets are

$$\{\emptyset, \mathscr{V}(\langle x, y \rangle), \mathscr{V}(\langle x - a, y - b \rangle), \mathscr{V}(I)\}$$

Let $U = X \setminus \mathscr{V}(\langle x, y \rangle)$ and $W = X \setminus \mathscr{V}(\langle x - a, y - b \rangle)$. The sheaf of ring $\mathscr{O}_X(X) = \mathbb{K}[x, y]/I$ can easily be localised, where $\mathscr{O}_X(U) = \mathbb{K}[x, y]/\langle x - a, y - b \rangle$ and $\mathscr{O}_X(W) = \mathbb{K}[x, y]/\langle x^i, y^j \rangle$.

❧ **Definition 11.** *Given two schemes X, Y, a* ***morphism of affine schemes*** *$f : X \to Y$ is a continuous map on the sets of primes, i.e. maps open sets to open sets, together with a pullback map on the underlying ringed structure, i.e. a map $f_* : \mathscr{O}_Y \to \mathscr{O}_X$, that respect the following*

$$\begin{aligned} X &\xrightarrow{f} Y \\ f^{-1}(U) &\mapsto U \\ \mathscr{O}_X &\xleftarrow{f_*} \mathscr{O}_Y \end{aligned}$$

where if $p \in U$ and if $g \in \mathscr{O}_Y(U)$ is such that $g(p) = 0$ then $f_*g(f^{-1}(p)) = 0$. A scheme morphism f is an isomorphism of schemes if and only if it is invertible [62, I.2].

Remark 2.1.2 (Correspondence). Generally, a scheme morphism $\operatorname{Spec} A \to \operatorname{Spec} B$ comes from a ring morphism $B \to A$. Thus, schemes $\operatorname{Spec} A$ and $\operatorname{Spec} B$ are isomorphic if and only if $A \cong B$ as rings. To compare the local structure at point $p \in \operatorname{Spec} A$ and $p' \in \operatorname{Spec} B$ look for morphisms between A_p and $B_{p'}$ [62, Theorem I-40].

Although it is a fun theory, we do not need most of it in this work, so we may simply summarise the important properties that bring nuance to our results. To enhance the connection with algebraic varieties, we may compare the two structures:

	affine varieties	**affine schemes**
intersection of plane curves	a point $p \in V(I)$	a maximal ideal in $\operatorname{Spec} R/I$
local structure (*ex.* : *multiplicity*)	forgotten	remembered
when $\mathbb{K} \neq \bar{\mathbb{K}}$	identical for some ideals *e.g.* $R = \mathbb{R}[x]$ $V(\langle x^2+1\rangle) = V(\langle 1\rangle)$	distinct for all ideals $\operatorname{Spec} \mathbb{R}[x]/\langle 1\rangle \neq \operatorname{Spec} \mathbb{R}[x]/\langle x^2+1\rangle$

Table 2.1: Vis-à-vis affine varieties and affine schemes.

As a topological set, there is a direct correspondence between the open sets of the algebraic variety of an ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$, $V(I)$, and $\mathbb{K} = \bar{\mathbb{K}}$, and $\operatorname{Spec} \mathbb{K}[x_1, \ldots, x_n]/I$

$$\{\text{open sets V(I)}\} \leftrightarrow \{\text{open sets } \operatorname{Spec} \mathbb{K}[x_1, \ldots, x_n]/I\};$$

$$\{\text{closed sets V(I)}\} \leftrightarrow \{\text{closed sets } \operatorname{Spec} \mathbb{K}[x_1, \ldots, x_n]/I\}.$$

If X is open (closed) in $\operatorname{Spec} \mathbb{K}[x_1, \ldots, x_n]/I$ then $\bigcup_{\mathfrak{p} \in X} V(\mathfrak{p})$ is open (closed) in $V(I)$.

To our specific interest, varieties *forget* the local structure, in contrast to schemes which preserve it. To illustrate this thought, in general $V(I) = V(\sqrt{I})$, in particular $V(\langle x \rangle) = V(\langle x^i \rangle)$. However, $\mathbb{C}[x]/\langle x \rangle$ and $\mathbb{C}[x]/\langle x^i \rangle$ for $i > 1$ fundamentally differ; note the existence of nilpotent elements in the latter. Let $X = \operatorname{Spec} \mathbb{C}[x]/\langle x \rangle$ and $Y = \operatorname{Spec} \mathbb{C}[x]/\langle x^i \rangle$, this translates to a nilpotent global section on the scheme Y (see Example 2.1.5 for $\mathscr{O}_X$ and $\mathscr{O}_Y$). In terms of regular functions, if we take a perspective from $\mathbb{C}[x]$, if $f, g \in \mathbb{C}[x]$ then $f = g$ as regular function on Y if f and g have the same canonical projection in $\mathbb{C}[x]/\langle x^i \rangle$, equivalently:

(**1**) $f(0) = g(0)$;

(**2**) $\frac{\partial^j f}{\partial x}(0) = \frac{\partial^j g}{\partial x}(0)$ for all $j < i$.

While $f = g$ in $\mathscr{O}_X(X)$ if and only if (**1**) holds. Based on this observation, the structure of the point is not identical and since there exists no ring isomorphism $\mathbb{C}[x]/\langle x \rangle \to \mathbb{C}[x]/\langle x^i \rangle \implies \operatorname{Spec} R/\langle x \rangle \ncong \operatorname{Spec} R/\langle x^i \rangle$ as schemes.

We can do a similar example to highlight that the multiplicity does not convey the full local structure.

❧ Example 2.1.7: Same multiplicity, different local structure

Let $X = \operatorname{Spec} \mathbb{C}[x, y]/\langle x^3, y \rangle$ and $Y = \operatorname{Spec} \mathbb{C}[x, y]/\langle x^2, xy, y^2 \rangle$ then the origin in X and Y are not locally isomorphic since $\mathscr{O}_X$ is a section of nilpotent order 3, *i.e.* $x^i \neq 0$ for $i < 3$ and $x^3 = 0$ in $\mathbb{C}[x, y]/\langle x^3, y \rangle$, which has no equivalent in $\mathscr{O}_Y$. Taking elements from $f = \sum_{i=0}^{n} \sum_{i=0}^{m} f_{i,j} x^i y^j, g = \sum_{i=0}^{n} \sum_{i=0}^{m} g_{i,j} x^i y^j \in \mathbb{C}[x, y]$, with their canonical projection, we see that f, g are equivalent regular functions on X if and only if

$$\underbrace{f_0}_{(\mathbf{1})} + \underbrace{f_{1,0} x + f_{2,0} x^2}_{(\mathbf{2})} = \underbrace{g_0}_{(\mathbf{1})} + \underbrace{g_{1,0} x + g_{2,0} x^2}_{(\mathbf{2})}$$

(1) $f(0,0) = g(0,0)$;

(2) $\frac{\partial^i f}{\partial x}(0,*) = \frac{\partial^i g}{\partial x}(0,*)$ for all $i < 3$.

while, f, g on Y, $f = g$ if and only if $f_0 + f_{1,0}x + f_{0,1}y = g_0 + g_{1,0}x + g_{0,1}y$, *i.e.*:

(1) $f(0,0) = g(0,0)$;

(2) $\frac{\partial^2 f}{\partial x}(0,*) = \frac{\partial^2 g}{\partial x}(0,*)$;

(3) $\frac{\partial^2 f}{\partial y}(*,0) = \frac{\partial^2 g}{\partial y}(*,0)$;

These represent the two $\langle x, y\rangle$-primary classes of multiplicity 3, *i.e.* the two possible local structures of multiplicity 3 up to isomorphism [122].

Later in this chapter present a similar non-monomial example, Example 2.1.18, and in Chapter 4, we discuss the concept of strata and their moduli space to study some ideals that share the same multiplicity.

Remark 2.1.3. Nilpotent sections of schemes, which witness multiplicity of intersection, can also be observed through algebraic varieties by choosing the proper ring extension[3].

❧ Example 2.1.8: Nilpotent elements and algebraic varieties

Take $i > j \in \mathbb{N}$:

over $\mathbb{Q}$	over $\bar{\mathbb{Q}}$	over $\mathbb{Q}[\alpha_l]$
$V_{\mathbb{Q}}(\langle x, y\rangle)$ $=$ $V_{\mathbb{Q}}(\langle x^i, y^j\rangle)$	$V_{\bar{\mathbb{Q}}}(\langle x, y\rangle)$ $=$ $V_{\bar{\mathbb{Q}}}(\langle x^i, y^j\rangle)$	$V_{\mathbb{Q}[\alpha_i]}(\langle x, y\rangle) = (0 + \langle \alpha_l^l\rangle, 0 + \langle \alpha_l^l\rangle)$ $\neq$ $V_{\mathbb{Q}[\alpha_l]}(\langle x^i, y^j\rangle) = (r_x + \langle \alpha_l^l\rangle, r_y + \langle \alpha_l^l\rangle)$

Table 2.2: Observations on ring extensions and affine varieties

where $V_R()$ defines the zero locus in $\mathbb{A}^2(R)$ and α_l is the residue class of $\mathbb{Q}[t]/\langle t^l\rangle$, $r_x \in \langle \alpha_l^{\min(i-l,1)}\rangle$ and $r_y \in \langle \alpha_l^{\min(j-l,1)}\rangle$.

[3] Interested readers may refer to [151] for a discussion relating schemes and algebraic varieties based on this observation.

So in a sense, as we want the local structure of the intersections, we aim to take a step towards schemes by finding a unique representation for each ideal I that describes the quotient ring R/I; this screams *Gröbner basis*. This will describe the ringed space and the maximal ideals of $\operatorname{Spec} R/I$, which we reduce to a problem of factorisation[4]. To focus on a given intersection p, we look at the localisation $\operatorname{Spec} R_p/I_p$.

2.1.2 Primary decomposition

As mentioned above, we are interested in the localisation of ideals $I \subseteq R$, with R a polynomial ring, and their quotient rings.

It is convenient, but mostly fun, to revisit some relations between a generating set of I, its primary decomposition and localisation at some points or hypersurfaces, as we will refer to them in Chapter 5 and Chapter 6. We are starting by reviewing some common reminders. Let R be a ring, if $\mathfrak{p} \subseteq R$ is a prime ideal, then $R_{\mathfrak{p}}$ is a local ring, i.e. a ring with a unique maximal ideal, where $\mathfrak{p}R_{\mathfrak{p}}$ is the maximal ideal [7, §3].

This leads to the useful property of *forgetting* primary components.

❧ **Definition 12.** *Let R be a commutative ring. An ideal $I \subseteq R$ is* ***primary*** *if I is proper and $ab \in I$ implies a or $b^i \in I$ for some $i \in \mathbb{N}$.*

In particular, $\sqrt{I}$ is prime if I is primary

❧ **Definition 13.** *Let I be an ideal, then $I = \bigcap_{Q \in \mathcal{Q}} Q$ is a* ***primary decomposition*** *of I if*

- *$\mathcal{Q}$ is finite ;*
- *the Q's are primary;*
- *it is irredundant, i.e., for all $Q, P \in \mathcal{Q}$, $\sqrt{Q} \neq \sqrt{P}$ and Q is not contained in the intersection of $\mathcal{Q} \setminus \{Q\}$*

The Q's are called the $\sqrt{Q}$-primary components of I [61, §3, proposition 3.9].

Over Noetherian rings, primary decomposition always exists for ideals $I \subseteq R$ and $\{\sqrt{Q} \mid Q \in \mathcal{Q}\}$ are the associated primes of the R/I as a R-module (Lasker–Noether, see [61, Theorem 3.10]), *i.e.* the primes in R that annihilate an element of R/I [7, proposition 4.5]. In particular, the set of radicals of the primary components is unique.

[4] We do not address the step of factorisation in this document, but efficient algorithms exist to factor polynomials, *e.g.*, [114] over $\mathbb{Q}$ and [13] over global fields.

Proposition 2.1.1. *When I is a zero-dimensional ideal, we have a bijection of sets*

$$\begin{array}{ccc} AP(\mathbb{K}[x_1,\ldots,x_n]/I) & \to & \operatorname{Spec}(\mathbb{K}[x_1,\ldots,x_n]/I) \\ P & \to & P\mathbb{K}[x_1,\ldots,x_n]/I, \end{array}$$

where AP is the set of primes of $\mathbb{K}[x_1,\ldots,x_n]$ associated with $\mathbb{K}[x_1,\ldots,x_n]/I$.

Proof. For zero-dimensional ideal I,

$$\begin{aligned} \operatorname{Spec}(\mathbb{K}[x_1,\ldots,x_n]/I) &= \left\{\text{maximal ideals of } \mathbb{K}[x_1,\ldots,x_n]/I\right\} \\ &= \left\{\mathfrak{m}\mathbb{K}[x_1,\ldots,x_n]/I \mid \mathfrak{m} \text{ maximal in } \mathbb{K}[x_1,\ldots,x_n] \text{ and } \mathfrak{m} \supseteq I\right\} \end{aligned}$$

(see Appendix C.2.1 for zero-dimensional ideals). Thus the bijection is a direct consequence of [7, proposition 4.6]. □

To our specific interest, if $I = \bigcap_{Q\in\mathcal{Q}} Q \subseteq R$ is a primary decomposition of I then for all $Q \in \mathcal{Q}$: $I_{\sqrt{Q}} = Q_{\sqrt{Q}}$ and $(R/I)_{\sqrt{Q}} \cong R/Q$ [61, §2.4, Theorem 3.10].

Remark 2.1.4 (Motivation for finding primary components). This previous observation supports the statement that if $I \subseteq \mathbb{K}[x_1,\ldots,x_n]$ is zero-dimensional, the local structure of a point $\mathfrak{p} \in \operatorname{Spec} R/I$, is purely defined by the corresponding primary component of I. If $I = \bigcap_{Q\in\mathcal{Q}} Q \subseteq R$ is a primary decomposition of I, then by Proposition 2.1.1 there exists $P \in \mathcal{Q}$ such that $\sqrt{P}R/I = \mathfrak{p}$. Thus $(R/I)_{\mathfrak{p}} \cong R_{\sqrt{P}}/I_{\sqrt{P}} \cong R/P$ hence $\operatorname{Spec}(R/I)_{\mathfrak{p}} \cong \operatorname{Spec} R/P$.

Thence, instead of describing $\operatorname{Spec}(R_{\mathfrak{p}}/I_{\mathfrak{p}})$, we chose to describe $\operatorname{Spec}(R/P)$, which gives us the (isomorphic) preimage of the quotient ring under the localisation. To achieve our ends, we choose the Gröbner basis of P, see summary infra, to describe the quotient.

Therefore, we preliminarily intend to extract the primary components of an intersection. To proceed, one could opt for a primary component decomposition algorithm. However, we may often get the same result by adding well-chosen generators when localising at an isolated point. Let $\mathbb{K} = \bar{\mathbb{K}}$, $R = \mathbb{K}[x_1,\ldots,x_n]$ with $I = \langle f_1,\ldots,f_t\rangle \subset R$ an ideal and let $\bigcap_{Q\in\mathcal{Q}} Q$ be a primary decomposition of I. Suppose $\sqrt{P}$ is maximal for some $P \in \mathcal{Q}$. By the Nullstellensatz $\sqrt{P} = \langle x_1 - \xi_1,\ldots,x_n - \xi_n\rangle$ for some $(\xi_1,\ldots,\xi_n) \in \mathbb{K}^n$; by definition of a radical ideal it follows that for all $i \in [1,\ldots,n]$, $(x_i - \xi_i)^{a_i} \in P$ for some $a_i \in \mathbb{N}^+$.

Lemma 2.1.1. *For $c_i \geq a_i$, $c_i \in \mathbb{N}$, then $\langle f_1,\ldots,f_t,(x_1-\xi_1)^{c_1},\ldots,(x_n-\xi_n)^{c_n}\rangle = P$*

Proof. Let $J = \langle (x_1 - \xi_1)^{c_1}, \ldots, (x_n - \xi_n)^{c_n} \rangle$, we may consider the union

$$I \cup J = \bigcap_{Q \in \mathcal{Q}} \Big(Q \cup J \Big).$$

Here for $Q \neq P$, $V(Q \cup J) = V(Q \cup P) = \emptyset$ thus by the Nullstellensatz, the union $Q \cup J = R$ and so

$$\begin{aligned} I \cup \langle (x_1 - \xi_1)^{c_1}, \ldots, (x_n - \xi_n)^{c_n} \rangle &= P \cup \langle (x_1 - \xi_1)^{c_1}, \ldots, (x_n - \xi_n)^{c_n} \rangle \\ &= P \end{aligned}$$

□

In some cases, we can reduce the number of generators, *i.e.*, if there exists a unique $p = (p_1, \ldots, p_n) \in V(I)$ such that $p_i = \xi_i$ for $i \in H$ where H is a subset of $\{1, \ldots, n\}$. We discuss the occurrence of this scenario in the plane and give an upper bound on the a_i's in Chapter 5. Nevertheless, for now, it is worth mentioning the obvious bound.

Lemma 2.1.2. *Let I be an ideal with a primary component P such that $(\xi_1, \ldots, \xi_n) = V(P)$, a point, and let δ be the multiplicity of I at $(\xi_1, \ldots, \xi_n)$, then there exists some $a_i \leq \delta$ such that the inclusion $(x_i - \xi_i)^{a_i} \in P$ holds.*

Proof. By definition, $\delta = \dim_{\mathbb{K}} R_{\sqrt{P}}/I_{\sqrt{P}}$, therefore $1, (x_i - \xi_i), (x_i - \xi_i)^2, \ldots (x_i - \xi_i)^{\delta}$ cannot be $\mathbb{K}$-linearly independent in $R_{\sqrt{P}}/I_{\sqrt{P}}$. Thus, there exists some $\mathbf{c} \in \mathbb{K}^{\delta+1} \setminus \mathbf{0}$ such that $\sum_{j=0}^{\delta} \mathbf{c}(j)(x_j - \xi_j)^j$ is equivalently 0 in $R_{\sqrt{P}}/I_{\sqrt{P}} \cong R/P$. Thus, $\sum_{j=0}^{\delta} \mathbf{c}(j)(x_i - \xi_i)^j$ lies in P . Let $l = min\{j \mid \mathbf{c}(j) \neq 0\}$ if $l = \delta$ then $(x_i - \xi_i)^{\delta} \in P$ which proves the statement. Otherwise, let $a_i \in \mathbb{N}^+$ be minimal such that $(x_i - \xi_i)^{a_i} \in P$, recall that the existence of a_i follows by Nullstellensatz and radically. Suppose $a_i > \delta$ then

$$(x_i - \xi_i)^{a_i - l - 1} \underbrace{\Big(\sum_{j=l}^{\delta} \mathbf{c}(j)(x_i - \xi_i)^j \Big)}_{\in P} = (x_i - \xi_i)^{a_i - 1}\mathbf{c}(l) + \underbrace{(x_i - \xi_i)^{a_i} \sum_{j=l+1}^{\delta} \mathbf{c}(j)(x_i - \xi_i)^{j-l}}_{\in P}$$

is in P. Thus $(x_i - \xi_i)^{a_i - 1} \in P$ which contradicts the minimality of a_i, hence $a_i \leq \delta$. □

2.1.3 Gröbner bases

Our method recovers a Gröbner basis of primary component(s), named I, with the motivation of describing $\mathbb{K}[x, y]/I$. Here, we summarise the foundations of Gröbner bases based on [45], [44] and [61]. We include exposition and applications to emphasise the depth of the results; experienced readers could skip this section. We review the underlying theory, the concepts, the definitions, the algorithms and the main theorems. More applications can be found in Appendix C where we also present the strong connections between Gröbner bases, syzygies, and algebraic geometry through some examples and by exposing some interesting connections with dimension theory.

Gröbner bases were originally introduced by a PhD student, Bruno Buchberger, who named the concept after his thesis advisor Wolfgang Gröbner. The bases were introduced for ideals in a polynomial ring over a field $\mathbb{K}[x_1, \ldots, x_n]$. However, when exposing the general concept in this section, we might as well review the generalization to free modules. This generality is notably useful for us for the syzygy of a generating set of modules. The symbol § is used whenever the purpose of a paragraph or an example is to add exposition and to help understand the amplitude of the representation without being critical to the rest of this document.

2.1.3.1 The Basics

Hereinafter, we let $\mathbb{K}$ be a field and $R = \mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial ring over $\mathbb{K}$ unless stated otherwise. We further let R^d be a free R-module generated by the standard basis $(e_1, \ldots, e_d)$. A **monomial** in R^d can be written as $e_i\mathbf{x}^a$ for some i, where $a = (a_1, \ldots, a_n) \in \mathbb{N}^n$ and $\mathbf{x}^a = \prod_{j=1}^n x_j^{a_j}$. Products and additions on elements of R^d are made components-wise. We now define the foundation of Gröbner bases: the monomial ordering.

❧ **Definition 14.** *A* ***monomial ordering*** *is an relation $\succ$ on monomials in R^d that satisfies:*

1. *the ordering is* ***total*** *in the sense that it is well-defined for each pair of monomials;*

2. *if a, b are monomials in R^d and c is a monomial in R, then $a \succ b \implies ac \succ bc$;*

Given a monomial ordering $\succ$, for any f in R^d, we define the initial term (or *leading term*) of f, $in(f)$, as the largest monomial in f with respect to $\succ$. In this section, we write $in_c(f)$ when we also include the corresponding coefficient. It is noteworthy that monomial orderings over R^d are not unique. Let $d = 1$, $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ in $\mathbb{N}^n$, then the following are monomial orderings:

❧ Example 2.1.9: Lexicographic order

Say $x_i \succ x_{i+1}$ [a] then $\mathbf{x}^a \succ_{lex} \mathbf{x}^b \iff$ for the smallest i such that $a_i \neq b_i$, $a_i - b_i$ is positive;

[a] For lexicographic order, most authors generally specify an order on the indeterminates. This is sufficient to define the relation between all the monomials for the lexicographic order.

❧ Example 2.1.10: §Graded Lexicographic Order

$\mathbf{x}^a \succ_{grlex} \mathbf{x}^b \iff \sum_{i=1}^n a_i > \sum_{i=1}^n b_i$ or $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$ and $\mathbf{x}^a \succ_{lex} \mathbf{x}^b$. In other words, we compare the degree first, and the lexicographic ordering breaks the ties.

❧ Example 2.1.11: §Reverse Graded Lexicographic Order

$\succ_{grevlex}$ is as above but in case of a tie $\mathbf{x}^a \succ_{lex} \mathbf{x}^b \implies \mathbf{x}^b \succ_{grevlex} \mathbf{x}^a$;

❧ Example 2.1.12: §Weighted Order

Based on some fixed $(w_1, \ldots, w_n) \in \mathbb{R}^n$, we define $\mathbf{x}^a \succ_{grlex} \mathbf{x}^b \iff$

$$(w_1, \ldots, w_n) \cdot (a_1, \ldots, a_n) > (w_1, \ldots, w_n) \cdot (b_1, \ldots, b_n)$$

where $\cdot$ is the dot products in $\mathbb{R}^n$. To be well-defined, this ordering requires that $(w_1, \ldots, w_n)$ is well-chosen in the sense that ties may not occur. An example is to consider $(w_1, \ldots, w_n)$ with $w_i/w_j \notin \mathbb{Q}$ for all i, j

❧ Example 2.1.13: §Block Order

Let $\mathbf{x}$ be split into c blocks $b_1, \ldots, b_c$ where $c \leq n$ and let $\succ_{b_i}$ be an ordering for each block. Then $\mathbf{x}^\alpha \succ_{block} \mathbf{x}^\beta \iff \mathbf{x}^\alpha \succ_{b_i} \mathbf{x}^\beta$ and $\mathbf{x}^\alpha$ and $\mathbf{x}^\beta$ are tied in $\succ_{b_j}$ for all $j < i$.

It is not a hard exercise to prove the above are monomial ordering; actual proofs can be found in [45][61]. Although monomial orderings are not unique, they all share useful proprieties, some of which we now review.

Lemma 2.1.3. *[61, §15.2] Any monomial order on R^d is Artinian, i.e. there exists no infinite descending chain of monomials $\cdots \prec m_{i-2} \prec m_{i-1} \prec m_i$ where the m_i's are monomials in R^d*

Lemma 2.1.4. *: Let $g, f_1, \ldots, f_t$ in R^d. Using a fixed arbitrary ordering $\succ$, then one may define some $r, q_1, \ldots, q_t \in R^d$ such that*

$$g = q_1 f_1 + \cdots + q_t f_t + r, \tag{2.1.1.1}$$

where $r = 0$ or $in(f_i) \nmid in(r)$ and $in(f_i q_i) \preceq in(g)$ for all f_i.

In other words, a monomial ordering may be used to define a division algorithm in R^d.

Proof. (inspired by [61, §15.3]) Fix $\succ$ and let $g, f_1, \ldots, f_t$ in R^d. If we let $q_i = 0\ \forall i$ and $r = g$ then

$$g = q_1 f_1 + \cdots + q_t f_t + r. \tag{2.1.1.2}$$

If $r = 0$ or $in(f_i) \nmid in(r)$ for any f_i, we are done. Otherwise, there exists some f_i such that $in(f_i) \mid r$. We replace $q_i \leftarrow q_i + in_c(r)/in_c(f_i)$ and let

$$r = r - (q_1 f_1 + \cdots + q_t f_t),$$

so r and q_i preserve the equality (2.1.1.2). Observe that $in(g) \succeq in(r) \succeq in(in_c(r)/in_c(f_i))$, thus $in(f_i q_i) \preceq in(g)$. Here, the initial term of r is eliminated; we may repeat the procedure with the newest initial term. We get that $in(r)$ is strictly decreasing at each step. By the lemma above, the procedure always terminates. □

In general, $r, q_1, \ldots, q_t$ in the above are not unique.

❧ **Definition 15.** *Given $f = (f_1, \ldots, f_d)$, $g = (g_1, \ldots, g_d)$ in R^d the least common multiple (lcm) and greatest common divisors (gcd) of f and g is defined component-wise.*

The above is well defined since R is a unique factorization domain (UFD). This last definition allows us to introduce S-polynomials.

❧ **Definition 16** (S-polynomial)**.** *Let $f, g \in R^d$. We define the* ***S-polynomial*** *of f and g to be $\sigma_{f,g} = \frac{lcm(in_c(f), in_c(g))}{in_c(f)} f - \frac{lcm(in_c(f), in_c(g))}{in_c(g)} g$.*

The terminology S-polynomial is more often used in the context of a polynomial ring, i.e. when $d = 1$. Note that if $in(f)$ and $in(g)$ are not supported on the same component e_i, element of the free basis, then $\sigma_{f,g} = 0$. The concept is analogous to syzygies which describe a linear relation between generators of a module.

❧ **Definition 17** (Syzygy)**.** *[61, §15.1, A3.9] Let $\mathcal{M}$ be an R-module, with R an arbitrary ring, generated by $g_1, \ldots, g_t$. The* ***first syzygy module*** *of $g_1, \ldots, g_t$ is the R-module $S_1(g_1, \ldots, g_t) \in R^t = \{(s_1, \ldots, s_t) \mid g_1 s_1 + \cdots + g_t s_t = 0\}$. Equivalently, $S_1(g_1, \ldots, g_t)$ is the module that satisfies the exact sequence*

$$0 \to S_1(g_1, \ldots, g_t) \to R^t \xrightarrow{\varphi_1} \mathcal{M} \to 0$$

where $(s_0, \ldots, s_t) \mapsto s_1\varepsilon_1 + \ldots + s_t\varepsilon_t$ for $\varepsilon_1 + \ldots + \varepsilon_t$ the standart basis of R^t and

$$\begin{array}{rccc} \varphi_1 : & R^t \cong \oplus_{i=1}^{t} \varepsilon_i R & \to & \mathcal{M} \\ & \varepsilon_i & \mapsto & g_i. \end{array}$$

Remark 2.1.5. The two parts of the definition of the first syzygy are truly equivalent: $S_1(g_1, \ldots, g_t) \to R^t$ is the inclusion $(s_1, \ldots, s_t) \mapsto s_1\varepsilon_1 + \cdots + s_t\varepsilon_t$. We observe that $(s_1, \ldots, s_t) \in ker(\varphi_1)$ if and only if $\varphi_1(s_1\varepsilon_1 + \cdots + s_t\varepsilon_t) = s_1 g_1 + \cdots + s_t g_t = 0$.

§ We defined what is called the first syzygy of a generating set of a module which gives a relationship between a free module and an arbitrary module. This is desirable as free modules are well-understood and possess many useful properties[5]. We note that $S_1(g_1, \ldots, g_t) = ker(\varphi)$ need not be free. If it is not, to describe the complete syzygy of a generating set of $\mathcal{M}$ one could define a long exact sequence

$$R^{t_{i+1}} \xrightarrow{\varphi_{i+1}} R^{t_i} \ldots \xrightarrow{\varphi_3} R^{t_2} \xrightarrow{\varphi_2} R^t \xrightarrow{\varphi_1} \mathcal{M} \to 0$$

where $\ker(\varphi_i)$ is the i^{th} syzygy of $g_1, \ldots, g_t$.

2.1.3.2 Gröbner Basis

By Hilbert's basis theorem: if $\mathbb{A}$ is Noetherian, then $\mathbb{A}[x_1, \ldots, x_n]$ is Noetherian. Since $\mathbb{K}$ is a field, it ensues that $R = \mathbb{K}[x_1, \ldots, x_n]$ is Noetherian. Thereby, any submodule of R^d is finitely generated as a module over R since R^d is a finitely generated free module. Hence, we will freely represent any submodule of R^d by an arbitrary finite generating set. Let $I \subseteq R$ be an ideal in R. For given monomial order $\succ$, we define the ideal of initial terms

$$\mathbf{In}(I) = \langle in(f) \mid f \in I \rangle,$$

[5] *e.g.*, they are flat and their Hilbert's series are easy to describe.

where $in()$ takes the initial term of its argument; we say that $\mathbf{In}(I)$ is the initial term ideal of I. A **monomial ideal** is characterized by $I = \mathbf{In}(I)$. For arbitrary submodule $\mathcal{M}$ of R^d with a fixed monomial ordering on R^d, we define $\mathbf{In}(\mathcal{M}) := \{in(m) \mid m \in \mathcal{M}\} = \bigoplus_i I_i e_i$ with each I_i a monomial ideal. $\mathcal{M}$ is a **monomial submodule** if $\mathcal{M} = \mathbf{In}(\mathcal{M})$.

❧ **Definition 18.** *A set $\mathcal{G} = \{g_1, \dots, g_s\}$ is a **Gröbner basis** of an R^d-submodule $\mathcal{M}$ with respect to a monomial ordering $\succ$ if it generates $\mathcal{M}$ and $\mathbf{In}(\mathcal{M}) = \langle in(g) \mid g \in \mathcal{G}\rangle$.*

When convenient, to keep track of the ordering between the generators, we sometimes use a vector to write Gröbner bases in this document. One of the main properties of Gröbner basis is that the divisions upon them (Lemma 2.1.4) have well-defined and unique remainders.

Theorem 2.1.2 (Macauley)**.** *[118][61, §,lemma 5.3] For any submodule $\mathcal{M}$ of $R^d = (\mathbb{K}[x_1, \dots, x_n])^d$ the monomials in $R \setminus \mathbf{In}(\mathcal{M})$, for any monomial ordering, form a $\mathbb{K}$-basis of the quotient ring $R^d/\mathcal{M}$*

Lemma 2.1.5. *Let $\mathcal{G}$ be a Gröbner basis of a R^d submodule for a monomial ordering $\succ$ and let $f \in R^d$ then the remainder of f upon the division with $\mathcal{G}$ is unique.*

Lemma 2.1.5 is a direct consequence of Theorem 2.1.2.

It is noteworthy that not all generating sets of submodules of R^d form a Gröbner basis (see example Example 2.1.14). Fortunately, a criterion exists to determine if a generating set forms a Gröbner basis.

Theorem 2.1.3. *(Buchberger's Criterion) Let $\mathcal{G} = (g_1, \dots, g_t)$ be a generating set of $\mathcal{M}$ as R^d-submodule and let r_{g_i,g_j} be the remainder of the division algorithm of σ_{g_i,g_j} by $g_1, \dots, g_t$. Then $\mathcal{G}$ forms a Gröbner basis of $\mathcal{M}$ if and only if $r_{g_i,g_j} = 0$ for all i, j.*

Proof. (inspired by [61, §15.4]) ($\Rightarrow$) Suppose $g_1, \dots, g_t$ form a Gröbner basis of $\mathcal{M}$. Then $\mathbf{In}(\mathcal{M}) = \langle in(g_1), \dots, in(g_t)\rangle$. Since the g_i are in $\mathcal{M}$, it follows that the S-polynomials σ_{g_i,g_j} are in $\mathcal{M}$. In the division algorithm,

$$r = \sigma_{g_i,g_j} - (q_1 g_1 + \cdots + q_t g_t);$$

hence r is also in $\mathcal{M}$. Therefore, at each step $in(r) \in \langle in(g_1), \dots, in(g_t)\rangle$ which implies divisibility. Thus the division algorithm only terminates when $r_{g_i,g_j} = 0$.

($\Leftarrow$) Suppose $r_{g_i,g_j} = 0$ for all i, j and suppose $g_1, \dots, g_t$ is not a Gröbner basis of $\mathcal{M}$; then there exists $m \in \mathcal{M}$ such that

$$in(m) \notin \langle in(g_1), \dots, in(g_t)\rangle.$$

Since $m \in \mathcal{M}$, there exist some $a_i \in R$ such that $m = a_1g_1 + \cdots + a_tg_t$.We choose m so that (in order)

1. $\alpha = \max(in(a_1g_1), \dots, in(a_1g_1))$ is minimal, and
2. $\mathcal{S} = \{i \mid in(a_ig_i) = \alpha\}$ is the smallest set possible

amongst all the elements that satisfy our assumption. Suppose α is supported by e_l. To get a cancellation of the initial terms $|\mathcal{S}| > 2$, otherwise we would have $\alpha = in(m) \in \langle in(g_1), \dots, in(g_t)\rangle$. Suppose $i, j \in \mathcal{S}$, then $in(a_ig_i) = in(a_jg_j) = \alpha$, thus $in(g_i) \mid \alpha$ and $in(g_j) \mid \alpha$; notably both are supported by e_l. It follows that $lcm(in(g_i), in(g_j)) \mid \alpha$ and $lcm(in(g_i), in(g_j))$ is also supported by e_l. Let $D(lcm(in(g_i), in(g_j))) = in_c(a_ig_i)$. By assumption $r_{g_i,g_j} = 0$, hence by Lemma 2.1.4 there exist some b_i's such that $\sigma_{g_i,g_j} = \sum_{i=1}^t b_ig_i$ with

$$in(b_ig_i) \leq in(\sigma_{g_i,g_j}) < lcm(g_i, g_j) \leq \alpha.$$

Using the two equality we have for σ_{g_i,g_j}

$$\begin{aligned}\tilde{m} &= m + D(\sigma_{g_i,g_j} - \sigma_{g_i,g_j}) \\ &= \sum_{k=1}^t a_kg_k + \sum_{k=1}^t Db_kg_k - D\frac{lcm(in(g_i), in(g_j))}{in(g_i)}g_i + D\frac{lcm(in(g_i), in(g_j))}{in(g_j)}g_j \\ &= \sum_{k=1}^t a'_kg_k\end{aligned}$$

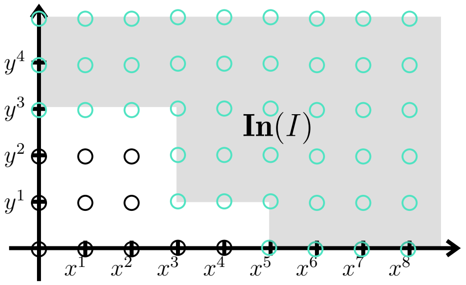
where

$$a'_k = \begin{cases} a_i + Db_i + D\frac{lcm(in(g_i),in(g_j))}{in(g_i)} & \text{if } k = i; \\ a_k + Db_j - D\frac{lcm(in(g_i),in(g_j))}{in(g_j)} & \text{if } k = j; \\ a_k + Db_k & \text{otherwise.}\end{cases}$$

By choice of D, $Db_kg_k \prec \alpha$ for all k and $in(D\frac{lcm(in(g_i),in(g_j))}{in(g_i)}g_i) = in(D\frac{lcm(in(g_i),in(g_j))}{in(g_i)}g_i) = \alpha$. Thus $in(a'_kg_k) \preceq max(in(a_kg_k), \alpha - 1)$. Moreover the initial terms of a_ig_i and $D\frac{lcm(in(g_i),in(g_j))}{in(g_i)}g_i$ cancel out thus we lose an occurrence of α contradicting the minimality of $|\mathcal{S}|$. Hence $g_1, \dots g_t$ form a Gröbner basis. □

❧ Example 2.1.14: Non-uniqueness of the generating set

Let $f_1 = y^3 + x^2y(x-1)$, $f_2 = x^3y + 43x^4, f_3 = x^5$. Consider $I \subseteq \mathbb{C}[x,y]$ generated by the following sets:



Gröbner basis	**Alternative**
$\langle\ f_1, f_2, f_3\ \rangle$	$\langle\ f_1,$ $f_2 + f_3 = x^3y + x^5 + 43x^4,$ $f_2 - f_3 = x^3y - x^5 + 43x^4\ \rangle$

We claim that the two sets generate the same ideal. Clearly $\langle f_1, f_2, f_3\rangle \supseteq \langle f_1, f_2+f_3, f_2-f_3\rangle$ and $\langle f_1, f_2, f_3\rangle \subseteq \langle f_1, f_2+f_3, f_2-f_3\rangle$ since $\frac{1}{2}(f_2-f_3)+\frac{1}{2}(f_2+f_3) = f_2$ and $\frac{1}{2}(f_2+f_3) - \frac{1}{2}(f_2-f_3) = f_3$. By Buchberger's criterion with the lexicographic ordering $y \succ x$, $\{f_1, f_2, f_3\}$ form **a** Gröbner basis.

S-Polynomials	**division in** (f_1, f_2, f_3)	r_{f_i,f_j}
$\sigma_{f_1,f_2} = -43x^4y^2 + x^6y - x^5y$	$0f_1 - (43xy)f_2 + (xy + (43^2-1)y)f_3$	0
$\sigma_{f_1,f_3} = x^8y - x^7y$	$0f_1 + 0f_2 + (x^3y + x^2y)f_3 + 0$	0
$\sigma_{f_2,f_3} = 43x^6$	$0f_1 + 0f_2 + 43xf_3 + 0$	0

Thus $\mathbf{In}(I) = \langle y^3, x^3y, x^5\rangle$. Since $x^5 \notin \langle in(f_1), in(f_2+f_3), in(f_2-f_3)\rangle = \langle y^3, yx^3\rangle$, the alternative generating set is not a Gröbner bases.

Remark 2.1.6. The figure on the left is called a monomials staircase. The grey area represents all the monomials in $\mathbf{In}(I)$ in lexicographic ordering $y \succ x$ while the staircase in white highlights a monomial basis of R/I seen as a $\mathbb{K}$-vector space (see Theorem 2.1.2).

In the previous example, we emphasized that we have **a** Gröbner basis of the module (the ideal) since Gröbner bases are not unique. When talking of ideals, some authors will often refer to **the** Gröbner basis of a module when it satisfies two conditions: being minimal and reduced. A basis is **minimal** if and only if the generators have no redundancies. We say g_i is redundant in a basis $g_1, \ldots, g_t$ if $g_i \in \langle g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_t\rangle$. A basis $g_1, \ldots, g_t$ where $in(g_i) \succ in(g_{i+1})$ is **reduced** when for all i the leading coefficient of g_i is 1 and for all $m \in \mathbf{In}(g_i)$, m does not appear in g_j for all $i \neq j$.

❧ **Example 2.1.15: Reduced and minimal**

In $\mathbb{C}[x,y]$ with lexicographic ordering induced by $y \succ x$ the following are all Gröbner bases of the same ideal

$\langle y^3 - x^2y - 43x^4, x^3y + 43x^4, x^5\rangle$ **is minimal and reduced**	$\langle y^3 + x^2y(x-1), x^3y + 43x^4, x^5\rangle$ **is minimal and not reduced**
$\langle y^3 - x^2y - 43x^4, x^3y + 43x^4, x^5y, x^5\rangle$ **is not minimal and reduced**	$\langle y^3 + x^2y(x-1), x^3y + 43x^4, x^5y, x^5\rangle$ **is not minimal and not reduced**

Theorem 2.1.4. *[45, §2.7, Theorem 5][61, Exercise 15.14] Let $\mathcal{M}$ an R^d submodule; then, for a fixed arbitrary monomial ordering, there exists a unique (up to reordering) Gröbner basis of $\mathcal{M}$ that is both minimal and reduced.*

It is easy to see that every Gröbner basis can be reduced and minimalized. Fix $\succ$, by Theorem 2.1.4 and by the division algorithm Lemma 2.1.5, Gröbner bases are a unique characterization an ideal and in quotient ring. We can also observe that strict containment is impossible for ideals having identical initial terms.

Lemma 2.1.6. *[61, §15.2, lemma 15.5] If $\mathbf{In}(\mathcal{M}_1) = \mathbf{In}(\mathcal{M}_2)$ have then $\mathcal{M}_1 \subset \mathcal{M}_2$ if and only if $\mathcal{M}_1 = \mathcal{M}_2$.*

2.1.3.3 Buchberger's Algorithm

One may not talk about Gröbner bases without presenting Buchberger's algorithm. The procedure allows one to find the Gröbner basis (and the syzygy of a generating set of a module, see Appendix C). It relies on the division algorithm and S-polynomials. Throughout, we consider some fixed monomial ordering $\succ$ on R^d.

Lemma 2.1.7. *There exists an algorithm that allows one to find a Gröbner basis of $\mathcal{M}$, a submodule of R^d, generated by $f_1, \dots, f_t$.*

Proof. (Buchberger) Given $f_1, \dots, f_t$ an arbitrary generating set, one may consider the following procedure:

Since the only elements we add to the set can be generated by $f_1, \dots, f_t$, we see that the resulting set generates the same module. By Buchberger's criterion, it is clear that the resulting set forms a Gröbner basis as the stopping condition is only reached when the remainder of all S-polynomials is 0. □

The only remaining issue is whether the above procedure always terminates, but we can show that the number of steps of Buchberger's algorithm is necessarily finite.

Algorithm 2.1.1 BUCHBERGER'S ALGORITHM($\mathcal{F}$)

```
INPUT: $\mathcal{F} = (f_1, \dots, f_t) \in R^d$
OUTPUT: $\mathcal{G} = (g_1, \dots, g_s)$ a Gröbner basis of the module generated by $f_1, \dots, f_t$
  $\mathcal{G} \leftarrow \mathcal{F}$
  Find $\sigma_{f,g}$ for all pairs $\{f, g\}$ in $\mathcal{G}$
  Add the remainder of $\sigma_{f,g}$ according to $\mathcal{G}$
  Remove the 0 from $\mathcal{G}$
  if $\mathcal{F} \neq \mathcal{G}$ then $\mathcal{G} \leftarrow$ Buchberger's algorithm on $\mathcal{G}$
  return $\mathcal{G}$
```

Lemma 2.1.8. *Buchberger's algorithm always terminates.*

Proof. Each iteration that does not terminate is characterized by adding some remainder r to the current generating set $\mathcal{G} = \{g_1 \dots, g_m\}$. By the division algorithm, the remainder satisfies the condition $in(g_i) \nmid in(r)$ for all i hence

$$\langle in(g_1) \dots, in(g_m), in(r) \rangle \supsetneq \langle in(g_1) \dots, in(g_m) \rangle.$$

By Lemma 2.1.3, the process is guaranteed to terminate. □

Note that most implementations of Buchberger's algorithm use additional steps to simplify $\mathcal{G}$ at each along the way to reduce the number of S-polynomials and remainders to compute and also to obtain a basis that is reduced and minimal.

Question 2.1.1. *By Lemma 2.1.7, there exists a way to find Gröbner basis, so why is this project relevant?*

While the algorithm terminates, the complexity cost is prohibitive. We discuss complexity at the end of this chapter.

2.1.3.4 Motivation

We now present the connection to the problem of describing the local structure of isolated points. Some additional Gröbner bases applications are presented in Appendix C. Henceforward we fix $d = 1$ and $R = \mathbb{K}[x_1, \dots, x_n]$ for $\mathbb{K}$ a field.

2.1.4 Connection to affine varieties

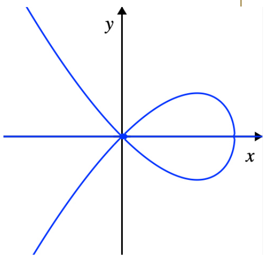
The variety of an ideal $I \subseteq R$, can easily be found from its Gröbner basis in **lexicographic** ordering. Assume the Gröbner bases are reduced minimal with $in(g_i) \succ in(g_{i+1})$. This basically boils down to a triangulation; the number of variables decreases as go *down* in the list of generators. Let $\mathcal{G} = (g_1, \ldots g_s)$ be the Gröbner basis of I. The idea is that g_s involves lesser variables than g_1. We can choose a root for the x_i's in g_s (which is a simpler problem - only one equation and fewer variables) we can inject it in g_{s-1} and repeat the process until we reach g_1. This finds elements of $V(I)$ provided that the leading coefficient polynomials of the $(g_1, \ldots, g_{s-j})$ do not vanish in the process [44, §2.1, Extension Theorem]. In the case of isolated points, the solutions of arbitrary polynomial equations are readable from the Gröbner basis – up to some factorization.

❧ **Example 2.1.16**

Taking the ideal from the earlier example

$$I = \langle y^3 + x^2y(x-1), x^3y + 43x^4, x^5 \ \rangle$$

the elements in the variety of I must be on the ordinate axis as they should lie in $V(x^5)$. Going up in the list of generators, injecting $x = 0$ in $x^3y + 43x^4$ does not restrict y, but when injecting in $y(y^2 + x^2(x-1))$ (variety drawn on the right) we get that $V(I) = \{(0,0)\}$.



This ordering reduces the problem of finding $V(I)$ to a problem of root finding in fewer variables. In the case where we have only isolated points in the affine variety (see below for dimension 0), the generators of the Gröbner basis form a triangular system on the x_i's, so the x_i can be solved one by one like the example above.

2.1.5 Connection with schemes

From Section 2.1.4 when R/I has Krull dimension 0 the lexicographic Gröbner bases efficiently describe $\operatorname{Spec} R/I$ as a set. More precisely, the points $p \in V(I)$, which can be found with the lexicographic ordering, have a correspondence with the ideals in $\operatorname{Spec} R/I$ which are all maximal when I is zero-dimensional (see Example 3.2.3 for zero-dimensional ideal), *i.e.* $\operatorname{Spec} R/I = \{I(p) \mid p \in V(I)\}$.

For a fixed monomial ordering $\succ$, Gröbner bases, when reduced and minimized, offer a unique representation of ideals (Theorem 2.1.4), which facilitates the comparison between ideals. Furthermore, by Macauley (Theorem 2.1.2), the monomials in $R^d \setminus \mathbf{In}(I)$, form a $\mathbb{K}$-basis of the quotient ring R/I. Let $p \in \operatorname{Spec} R/I$ and Q be the primary $I(p)$-primary component of I. The number of monomial in the basis of $R \setminus \mathbf{In}Q$, *i.e.* $\dim_{\mathbb{K}}\big(R_p/I_p\big)$, corresponds to the multiplicity at p. In more generality, through the division algorithm, a Gröbner basis of I accurately describes R/I; recall that this corresponds to the ring of regular functions on $\operatorname{Spec} R/I$. When focusing on a point $p \in\in \operatorname{Spec} R/I$, the Gröbner basis of the p-primary component of I, which describes the quotient ring $R/Q \cong R_p/I_p$, gives the ring of regular functions localized at p.

As stated above $\mathbf{In}(I)$ is sufficient to describe the quotient ring R/I as vector space, which can raise the following question.

Question 2.1.2. *Why do we wish to find the exact Gröbner basis of I through an $\mathfrak{m}$-adic expansion while the canonical image of the basis in $\mathbb{A}/\mathfrak{m}[x, y]$, with $\mathbb{A} \subset \mathbb{K}$ a domain and $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal, could be sufficient for finding $\mathbf{In}(I)$?*

Firstly $V(I)$ is generally not equal to $V(\mathbf{In}(I))$.

❧ **Example 2.1.17: $V(\mathbf{In}(I)) \neq V(I)$**

For example, take the lexicographic ordering induced by $(y \succ x)$ and let $I = \langle y^3 + 1, x^2 + 1\rangle \subseteq \mathbb{C}[x, y]$ and $I' = \langle y^3 + 1, x^2\rangle \subseteq \mathbb{C}[x, y]$, then $V(I) \neq V(I')$.

Moreover, getting the monomial ideal is not enough to describe the local structure.

❧ **Example 2.1.18**

For example, let $I = \langle y^3, x^2\rangle \subseteq \mathbb{C}[x, y]$ and $I' = \langle y^3 + x, x^2\rangle \subseteq \mathbb{C}[x, y]$, then $V(I) = V(I') = \{(0, 0)\}$. Both generating sets are lexicographic Gröbner bases with respect to $y \succ x$ and $\mathbf{In}(I') = I$. However, the minimal polynomial of y as degree 6 in $\mathbb{C}[x, y]/I$ and $\mathbb{C}[x, y]/I'$ contains no element for which the minimal polynomial is greater than 4. Thus, the ring of regular functions of $\operatorname{Spec} \mathbb{C}[x, y]/\langle y^3, x^2\rangle$ and $\operatorname{Spec} \mathbb{C}[x, y]/\langle y^3 + x, x^2\rangle$ are not isomorphic. Thus the local structure at the origin differs.

Furthermore, many algebraic applications and properties [6], some that we list in Appendix C, require the Gröbner basis and the initial terms alone just do not do the trick.

2.1.5.0.1 In short Gröbner basis is a simple description of modules. They require the definition of a monomial ordering, which notably allows us to perform division in the free modules. Gröbner bases are not unique and may be found by using Buchberger's algorithm. More algorithms are reviewed in the last section of this chapter. They simplify many questions from algebra and geometry: to our main interest, the description of schemes of zero-dimensional ideals.

2.1.6 Limits

We are primarily interested in using completion, a form of limit, for Gröbner bases algorithms to reduce the arithmetic complexity (see details in Chapter 4). The general concept of limits arises from category theory, and the idea is rather simple:

❧ **Definition 19.** *$\mathcal{C}$ is a **category** is a set/class of objects with a set/class of homomorphisms on the objects, including automorphisms, that satisfies composition [138, §1].*

❧ **Definition 20.** *Given a category $\mathcal{C}$, the **limit of** $\mathcal{C}$ ($\varprojlim \mathcal{C}$) is an object L in the category for which there exists a morphism to all the other objects in the category [138, §3].*

We are only interested in $\mathfrak{m}$-adic completion of a ring R with respect to a maximal ideal $\mathfrak{m} \subseteq R$ so we dedicate this section to them. A more general introduction to limits can be found in [138, §3].

❧ **Definition 21.** *Consider the category that contains the polynomial rings in n variables over $\mathbb{A}/\mathfrak{m}, \mathbb{A}/\mathfrak{m}^2, \mathbb{A}/\mathfrak{m}^3, \dots$ where $\mathbb{A}$ is a domain and $\mathfrak{m} \subseteq \mathbb{A}$ is a maximal ideal. Here, $\mathbb{A} \supseteq \mathfrak{m} \supseteq \mathfrak{m}^2 \supseteq \mathfrak{m}^3 \supseteq \dots$ is called the $\mathfrak{m}$-adic filtration [61, §7.1], and it admits the existence of canonical homomorphism between objects $\mathbb{A}/\mathfrak{m}^i \twoheadrightarrow \mathbb{A}/\mathfrak{m}^j$ for $i \geq j$. The $\mathfrak{m}$**-adic completion** of $\mathbb{A}$, denoted $\hat{\mathbb{A}}_{\mathfrak{m}}$ corresponds to the limit below:*

$$\hat{\mathbb{A}}_{\mathfrak{m}} := \varprojlim \mathbb{A}/\mathfrak{m}^i = \{(a_1, a_2, \dots) \mid a_i \in \mathbb{A}/\mathfrak{m}^i \textit{ and } a_i \equiv a_j \mod \mathfrak{m}^j \quad \forall i \geq j\} \hookrightarrow \prod_{i \geq 0} \mathbb{A}/\mathfrak{m}^i.$$

[6] *e.g.* the following properties are not preserved under Gröbner degeneration: primary, radical [41], Cohen-Macaulay ...

In other words, the limit is the object above them all for which we get a morphism to $\mathbb{A}/\mathfrak{m}^i$ for all i. Here, $\hat{\mathbb{A}}_\mathfrak{m}$ is a local ring.

❧ **Definition 22.** *Let $R = \mathbb{A}[x_1, \dots, x_n]$, we define the* $\mathfrak{m}$***-adic completion of a polynomial ring*** *R, $\hat{R}_\mathfrak{m}$, as the polynomial ring with coefficients in $\hat{\mathbb{A}}_\mathfrak{m}$.*

❧ Example 2.1.19: Integers p-adic completion

Let $\mathbb{A} = \mathbb{Z}$ and p be a prime number, then the completion based on the $\langle p \rangle$-adic filtration is $\hat{\mathbb{Z}}_{\langle p \rangle} = \mathbb{Z}_p$ are the p-adic numbers.

❧ Example 2.1.20: Ring extension

Let $\mathbb{A} = \mathbb{Z}[\alpha_1, \dots, \alpha_m]$ where the α_i's are algebraic[a] and let p be a prime number, then $\hat{\mathbb{A}}_{\langle p \rangle} = \mathbb{Z}_p \times \mathbb{Z}_p \alpha_1 \times \cdots \times \mathbb{Z}_p \alpha_m$.

[a] Recall: α is algebraic over a ring R if and only if there exists $f(x) \in R[x]$, $f \neq 0$ with $f(\alpha) = 0$.

Remark 2.1.7. When using the change of coordinates from Proposition 3.4.1 on a generating set in $\mathbb{Z}[x, y]$ we would get generators in $\mathbb{Z}[\alpha_1, \dots, \alpha_m][x, y]$ (thus the completion would be Example 2.1.20 instead of Example 2.1.19).

❧ Example 2.1.21

Let $\mathbb{A} = k[t_1, \dots, t_n]$ where k is a field and let $\mathfrak{m} = \langle t_1, \dots, t_n \rangle \subseteq \mathbb{A}$, then $\hat{\mathbb{A}}_\mathfrak{m} \cong k[\![t_1, \dots, t_n]\!]$, *i.e.* the ring of formal power series.

In this document, a limit *construction* of an element $a = (a_1, a_2, \dots) \in \hat{\mathbb{A}}_\mathfrak{m}$, is the iterative process of finding a by recovering the a_i's

$$a_1 \to (a_1, \dots, a_{i_1}) \to (a_1, \dots, a_{i_2}) \to \dots$$

for $i_1 < i_2 < \dots$.

Limit construction can be studied using algebraic varieties via Hensel's Lemma and Newton's method.

Theorem 2.1.5 (Hensel's Lemma/Newton's method). *[63][§7, Theorem 7.3, Exercise 7.26]. Let $\mathbb{A}$ be a domain and let $(f_1, \dots, f_n) \in \mathbb{A}[x_1, ..., x_n]^n$. If $\mathfrak{m} \subseteq \mathbb{A}$ is a maximal ideal and $p_1 \in (\mathbb{A}/\mathfrak{m})^n$ is such that $f_1(p_1) \equiv \cdots \equiv f_n(p_1) \equiv 0 \mod \mathfrak{m}$ then there exists $p = (p_1, p_2, \dots) \in V(\langle f_1, \dots, f_n \rangle) \cap (\hat{\mathbb{A}}_{\mathfrak{m}})^n$ with $p_i \in (\mathbb{A}/\mathfrak{m}^i)^n$ which satisfy the relation*

$$p_{2^j} \equiv \left(\overline{a_{2^{j-1}}} - J_{\langle f_1, \dots, f_n \rangle}(\overline{p_{2^{j-1}}})^{-1} \begin{bmatrix} f_1(\overline{p_{2^{j-1}}}) \\ \vdots \\ f_n(\overline{p_{2^{j-1}}}) \end{bmatrix} \right) \mod \mathfrak{m}^{2^j}$$

for all $j \in \mathbb{N}^{>1}$, where $J := \left[\nabla f_1, \dots, \nabla f_n \right]$ is the Jacobian of $\langle f_1, \dots, f_n \rangle$ and $\overline{p_{2^{j-1}}}$ is the lift in $\mathbb{A}/\mathfrak{m}^{2^j}$. The above converges if J is full rank at p and p_1.

Newton's method is an iterative technique that applies the relation of Theorem 2.1.5 to find p_i's (element in $(\mathbb{A}/\mathfrak{m}^i)^n$) for indexes that grow quadratically; this is the meaning of **quadratic convergence**. An iteration is commonly named *lifting*.

Remark 2.1.8. If $I = \langle f_1, \dots, f_n \rangle$ radical, the method can only converge if $p \in V(I)$ is **smooth**.

Remark 2.1.9. Like used Example 2.1.24, Newton's method works to find zeros of functions $f_i : \mathbb{A}^n \to \mathbb{A}$ for which the Jacobian is well defined, *e.g.* rational functions, in which case we do not talk about algebraic varieties. However, in this thesis, we only use Newton's method with f_i that can be seen as polynomials in $\mathbb{A}[x_1, \dots, x_n]$.

If $I = \langle f_1, \dots, f_m \rangle \subseteq \mathbb{K}[x_1, \dots, x_n]$ for $m > n$, *i.e.* J is not square, we can replace $J(p_j)^{-1}$ by $(J^\top J)^{-1} J^\top(p_j)$ or we can drop linearly dependant columns.

❧ **Example 2.1.22: Lift over a principal ideal domain**

Given a prime number p and the group embedding $\pi_{i+1} : \mathbb{Z}/p^{i+1}\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p^i\mathbb{Z}$, a **lifting in** $\hat{\mathbb{Z}}_{\langle p \rangle}$ constructs a limit with respect to the maps π_i, such that

$$\pi_i(a_i) = a_{i-1}$$

where $a_i := i^{th}$ lift. It can be seen as an operation to *increase the precision* by finding the image by going upstream in the sequence

$$\mathbb{Z}/p\mathbb{Z} \twoheadleftarrow \mathbb{Z}/p^2\mathbb{Z} \twoheadleftarrow \cdots \twoheadleftarrow \mathbb{Z}/p^{2^j}\mathbb{Z} \twoheadleftarrow \mathbb{Z}/p^{2^{j+1}}\mathbb{Z} \twoheadleftarrow \cdots$$

Here, the full chain would give the image in the completion, which is the p-adic numbers: $\mathbb{Z}_p$.

❧ Example 2.1.23: Lift over a function field

Let $R = k[t_1, \dots, t_m]$ where k is a field and let $\mathfrak{m} \subseteq k[t_1, \dots, t_m]$ be a maximal ideal. Consider the group surjections $\pi_{i+1} : k[t_1, \dots, t_m]/\mathfrak{m}^{i+1} \twoheadrightarrow k[t_1, \dots, t_m]/\mathfrak{m}^i$, a **lifting in** $\widehat{k[t_1, \dots, t_m]}_{\mathfrak{m}}$ constructs a limit with respect to the maps π_i, such that

$$\pi_i(a_i) = a_{i-1}$$

where $a_i := i^{th}$ lift. Again an iteration corresponds to finding a precursory image to go upstream in a sequence of extension

$$R/\mathfrak{m} \twoheadleftarrow R/\mathfrak{m}^2 \twoheadleftarrow \cdots \twoheadleftarrow R/\mathfrak{m}^{2^j} \twoheadleftarrow R/\mathfrak{m}^{2^{j+1}} \twoheadleftarrow \cdots$$

Here, the completion is isomorphic to the ring of formal power series $k[\![t_1, \dots, t_m]\!]$.

Since we get a quadratic convergence, it is convenient to focus only on the category that contains

$$\mathbb{A}/\mathfrak{m}, \mathbb{A}/\mathfrak{m}^2, \mathbb{A}/\mathfrak{m}^{2^2}, \dots$$

and to write the completion as

$$\hat{\mathbb{A}}_{\mathfrak{m}} = \{(a_0, a_1, \dots) \mid a_i \in \mathbb{A}/\mathfrak{m}^{2^i} \text{ and } a_i \equiv a_j \mod \mathfrak{m}^{2^j} \quad \forall i \geq j\}.$$

To use Newton's method to find an element p in $\mathfrak{m}$-adic completion of a ring R, one must define an ideal $I \subseteq R$ such that $p \in V(I) \subseteq \mathbb{A}^1(\hat{R}_{\mathfrak{m}})$ is smooth and isolated.

❧ Example 2.1.24: Inverses

Let $\mathbb{A}$ be a ring and $\mathfrak{m} \subseteq \mathbb{A}$ be a maximal ideal. Further let $a \in \mathbb{A} \setminus \mathfrak{m}$ and let $\hat{a} = (a_0, a_1, \dots)$ be the canonical image of a in $\hat{\mathbb{A}}_{\mathfrak{m}}$ with $a_i \in \mathbb{A}/\mathfrak{m}^{2^i}$ and $a_i \equiv a \mod \mathfrak{m}^{2^i}$. Retrieving the inverse of $\hat{a}$ is equivalent to finding a root of $f(x) = \frac{1}{x} - \hat{a}$. A first approximation for $\hat{a}^{-1} = (\xi_0, \xi_1, \dots)$ is $\xi_0 = a_0^{-1}$ (well defined since $a_0 \in \mathbb{A}/\mathfrak{m}$ which is a field), then

$$\begin{aligned} \xi_{i+1} &= \overline{\xi_i} + \frac{f(\overline{\xi_i})}{f'(\overline{\xi_i})} \in \mathbb{A}/\mathfrak{m}^{2^{i+1}} \\ &= 2\overline{\xi_i} - a\overline{\xi_i}^2 \end{aligned}$$

where $\overline{\xi_i}$ is a lift of ξ_i in $\mathbb{A}/\mathfrak{m}^{2^{i+1}}$. Let π_i be a lift from $\mathbb{A}/\mathfrak{m}^{2^i}$ to $\mathbb{A}$, we note that by construction, $a\pi_i(\xi_i) \equiv 1 \mod \mathfrak{m}^{2^i}$: firstly $\pi_0(a_0^{-1})a \equiv a_0^{-1}a_0 \equiv 1 \mod \mathfrak{m}$. Assume $a\pi_i(\xi_i) \equiv 1 \mod \mathfrak{m}^{2^i}$, then $a\pi_i(\xi_i) = 1 + r_i$, where $r_i \in \mathfrak{m}^{2^i}$. It follows that

$$a\pi_i(\xi_{i+1}) = a(2\pi_i(\xi_i) - a\pi_i(\xi_i)^2) = 1-(a\pi_i(\xi_i)-1)^2 = 1-(r_i)^2 = 1 + r_{i+1}$$

with $r_{i+1} \in \mathfrak{m}^{2^{i+1}}$.

Example 2.1.25: Approximation of a zero in a principal ideal ring

Let $\mathfrak{m} = \langle t \rangle \subseteq \mathbb{C}[t]$ and $f(x) = (t+1)x^2 - 1$. We may use Newton's method to find an element of $V(\langle (t+1)x^2 - 1 \rangle) = (\pm\sqrt{\frac{1}{(t+1)}})$ in the $\mathbb{C}[\![t]\!]$. Since $(1)(t+1) \equiv 1 \mod \langle t \rangle$, it follows that $f(1) \equiv 0 \mod \langle t \rangle$. Thus, we fix $\xi_0 = [1]$. Then for $\xi_i = \sqrt{\frac{1}{(t+1)}} \mod t^{2^i}$

$$\begin{aligned} \xi_{i+1} &\equiv \overline{\xi_i} - \frac{f(\overline{\xi_i})}{f'(\overline{\xi_i})} \mod t^{2^{i+1}} \\ &\equiv \overline{\xi_i} - \frac{(t+1)\overline{\xi_i}^2 - 1}{2(t+1)\overline{\xi_i}} \mod t^{2^{i+1}} \end{aligned}$$

where $\overline{\xi_i}$ is a lift of ξ_i in $\mathbb{C}[t]/\langle t^{2^{i+1}} \rangle$

at iteration 1: $f(1) = t$, $f'(1) = 2(t+1)$ and $\xi_1 \equiv 1 - t/(2t+2) \equiv 1 - t(\frac{-t}{2} + \frac{1}{2}) \mod t^2$ (the inverse $\frac{1}{2t+2} \equiv \frac{-t}{2} + \frac{1}{2} \mod t^2$ can be found using the above). Check that $\xi_1 \equiv \frac{-t}{2} + 1 \mod t^2$ and $f(\xi_1) \equiv 0 \mod t^2$.

at iteration 2: $f(\xi_1) \equiv \frac{1}{4}t^3 - \frac{3}{4}t^2 \mod t^4$, $f'(\xi_1) \equiv -t^2 + t + 2$ and $\xi_2 \equiv \frac{-t}{2} + 1 + \frac{\frac{1}{4}t^3 - \frac{3}{4}t^2}{-t^2+t+2} \equiv \frac{-t}{2} + 1 - (\frac{1}{4}t^3 - \frac{3}{4}t^2)(\frac{1}{16}(-5t^3 + 6t^2 - 4t + 8)) \equiv -\frac{5t^3}{16} + \frac{3t^2}{8} - \frac{t}{2} + 1 \mod t^4$ (the inverse $\frac{1}{-t^2+t+2} \equiv \frac{1}{16}(-5t^3 + 6t^2 - 4t + 8) \mod t^4$ can be found using the above). Here when $\xi_2 \equiv -\frac{5t^3}{16} + \frac{3t^2}{8} - \frac{t}{2} + 1 \mod t^4$, we get $f(\xi_2) \equiv 0 \mod t^4$.

and so on.

A last question arises: one can simply think of p-adic numbers to recall that elements in the completion can correspond to an infinite sequence. Hence, are we guaranteed that an algorithm based on the Newton iteration in Theorem 4.1.1 can eventually terminate?

Luckily the answer is yes in some cases. Let $\mathbb{A}, \mathbb{K}, \mathfrak{m}, I$ and $\mathcal{G}$ be as defined in Theorem 4.1.1, and further assume $\mathbb{K}$ is *large enough* and with an Archimedean or ultrametric valuation. Here $\mathcal{G}$, the minimal reduced lexicographic Gröbner basis of the ideal I, is in $\mathbb{K}[x, y]$. Using Hermite normal forms, we show in Chapter 5 that we may define a bound for the coefficients of $\mathcal{G}$. We also show that for a generic maximal ideal $\mathfrak{m} \in I$, $\mathcal{G} \subset \mathbb{A}_{\mathfrak{m}}[x, y] \subset \mathbb{K}[x, y]$, where $\mathbb{A}_{\mathfrak{m}}$ is the localisation of $\mathbb{A}$ at $\mathfrak{m}$. Therefore, the minimal reduced Gröbner basis of the ideal I in Theorem 4.1.1 lives in $\left\{\frac{a}{b} \in \mathbb{K} \mid |a| < |h|, |b| < |h|, b \notin \mathfrak{m}\right\}$ for a well chosen $\mathfrak{m}$ and a known $h \in \mathbb{A}$. It follows that if there exists a fraction reconstruction algorithm with respect to the $\mathfrak{m}$-adic filtration, one can find a Gröbner basis of I in $\mathbb{K}[x, y]$ that is exact after some iterations.

❧ **Definition 23.** *Let $\mathbb{A}$ be a domain, $\mathfrak{m} \subseteq \mathbb{A}$ be a maximal ideal and $\mathbb{K}$ be the fraction field of $\mathbb{A}$ with an Archimedean or ultrametric valuation $|\,|$. In this document, we say there exists a* ***fraction reconstruction algorithm with respect to the $\mathfrak{m}$-adic filtration*** *if for an injection*

$$f_i : \left\{\tfrac{a}{b} \in \mathbb{K} \mid |a| < |h|, |b| < |h|, b \notin \mathfrak{m}\right\} \hookrightarrow \mathbb{A}/\mathfrak{m}^i,$$

for some $h \in \mathbb{A}$ and $i \in \mathbb{N}$, there exists an algorithm that terminates to recover $f_i^{-1}(a)$ for all $a \in \mathbb{A}/\mathfrak{m}^i$.

In the case of some rings, fraction reconstruction algorithms exist for any maximal ideal. This is the case of $\mathbb{A} = \mathbb{Z}$ ($\mathbb{K} = \mathbb{Q}$) and $\mathbb{A} = k[t]$ ($\mathbb{K} = k(t)$, the function field of the affine line for a field k) [20, §7.1] [73, §5.7, Theorem 1.16, Corollary 1.7] and $\mathbb{A} = k[t_1, \ldots, t_n]$ ($\mathbb{K} = k(t_1, \ldots, t_n)$) [142].

2.1.7 Moduli Space of Strata

In order to use Newton's method to find the Gröbner basis of an ideal I, and by this mean characterising the local structure of intersections, we must define an alternative ideal $\mathscr{I}$ with $p \in V(\mathscr{I})$ such that

- the Gröbner basis of an ideal I can be found from p;
- p is smooth to preserve the convergence.

The key idea of Chapter 4 is to use the property that if I, J are two ideals over the same ring, then $(I+J)/J = 0$ (*i.e.* $a \mod J \equiv 0$ for all $a \in I$) if and only if $I \subseteq J$. Since it would not be practical to test every possible ideal separately, we rely on a (small) set of ideals that contains I: the stratum.

❧ **Definition 24.** *Let J a monomial ideal in a polynomial ring R, the* ***Gröbner cell*** *(or* ***stratum****) corresponding to J is defined as the set of ideals*

$$\mathcal{C}(J) := \{J \subseteq R \mid \mathbf{In}(I) = J\}.$$

If $\boldsymbol{E}$ is the Gröbner basis of I, we equivalently write $\mathcal{C}(I)$ or $\mathcal{C}(\boldsymbol{E})$ [33, 32, 95, 69]

Chapter 4 shows how the moduli space of a stratum can be used to define the ideal to base our Newton iteration.

❧ **Definition 25** (informal)**.** *Geometry objects (such as sets of ideals sharing some properties, genus g curves, vectors bundles, etc.) and their homomorphisms/isomorphisms and automorphisms can sometimes be studied via the space associated with those objects. The space of a set of objects, which generally corresponds to a parametrisation of the set, is named* ***moduli space*** *when it exists [133].*

In general, moduli space are *stacks*; however, in the specific case of strata, the moduli spaces turn out to be affine spaces[7] [131, 139, 113] which gives us a nice topology to work on!

For bivariate ideals of Krull dimension 0, there exists a parametrisation of the syzygies based on Hilbert-Burch matrices introduced by Conca *&* Valla [40] from which emerge explicit map between an ideal in the stratum and a point in the corresponding moduli space. First, it is interesting to recall Lazard's structure theorem for bivariate lexicographic Gröbner bases [109].

Theorem 2.1.6 (Lazard's structure theorem)**.** *[109, Theorem 2] If $I \subseteq \mathbb{K}[x,y]$ is an ideal, then there exists*

- $D_0, \ldots, D_s \in \mathbb{K}[x]$
- $F, G_0, \ldots, G_s \in \mathbb{K}[x,y]$ *where the G_i's are monic with respect to y in the ideal*

  $$\langle G_{i-1}, D_{i+2}G_{i+2},\ \ldots,\ D_{i+2}\cdots D_{s-1}G_s,\ D_{i+2}\cdots D_s\rangle$$

  and $G_s = 1$

[7] For this reason, we do not review the general definition of stacks in this document, interested readers may refer to [133]. Nonetheless, one should be aware that general moduli spaces are not always schemes.

subject to the constraint

$$\{FG_i \prod_{j=0}^{i} D_j \mid i \in [0, s]\}$$

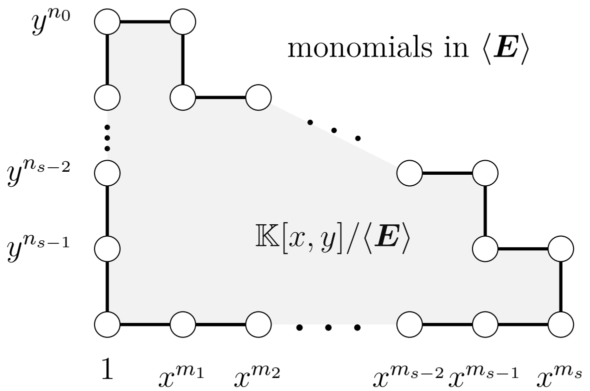
form the lexicographic ($y \succ x$) Gröbner basis of I.

Conca *&* Valla's map only applies to bivariate zero-dimensional ideals. In particular, we consider monomial ideals with bases of the form

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s}),$$

$n_i > n_{i+1}$ and $m_i < m_{i+1}$, whose Gröbner cells are stable under intersection with the set of zero-dimensional ideals

$$\mathcal{C}(\boldsymbol{E}) \subseteq \{I \subseteq \mathbb{K}[x, y] \mid \mathbf{In}(I) = \langle \boldsymbol{E} \rangle, |V(I)| < \infty\}.$$



y^{n_0}
monomials in $\langle \boldsymbol{E} \rangle$
$y^{n_{s-2}}$
$y^{n_{s-1}}$
$\mathbb{K}[x,y]/\langle \boldsymbol{E} \rangle$
1
x^{m_1}
x^{m_2}
$x^{m_{s-2}}$
$x^{m_{s-1}}$
x^{m_s}

Figure 2.1: Monomials staircase of $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$

Let Λ be a set of variables, for $i = 1, \dots, n_s + 1$, define g_i to be the n_0-minor in $\mathbb{K}[\Lambda][x, y]$ obtained when removing the i^{th} row of the parametric Hilbert-Burch matrix of $\mathcal{C}(\boldsymbol{E})$ given by:

$$\boldsymbol{H} = \begin{bmatrix} -x^{d_1} & & & \\ y & -x^{d_2} & & \\ & \ddots & \ddots & \\ & & y & -x^{d_{n_0}} \\ & & & y \end{bmatrix} + \begin{bmatrix} -n_{1,1} & & \\ & & \\ \vdots & \ddots & \\ & & -n_{n_0,n_0} \\ -n_{n_0+1,1} & \cdots & -n_{n_0+1,n_0} \end{bmatrix},$$

where $n_{i,j} = \sum_{0\leq l<d_i} \Lambda_{l,i,j}x^l$ with

$$d_i = \begin{cases} m_r - m_{r-1} & \text{if } n_r = i \\ 0 & \text{otherwise .} \end{cases}$$

(width of the monomial step at y^r if any)

and the $\Lambda_{l,i,j}$'s are variables in Λ. Let $N = |\Lambda|$; there exists a bijection between $\mathbb{A}^N(\mathbb{K})$ and $\mathcal{C}(\boldsymbol{E})$ through the generating set given by the g_i: when we evaluate Λ at a point in $\mathbb{A}^N(\mathbb{K})$, we obtain Gröbner basis of an ideal in $\mathcal{C}(\boldsymbol{E})$. Their result follows for the structure of the syzygies of the ideal in $\mathcal{C}(\boldsymbol{E})$. The columns of $\boldsymbol{H}$ generate a Gröbner basis of the syzygy module: they rely on Schreyer Theorem (see Proposition C.1.2, [61, Theorem 15.10]) to prove they form a generating set of the syzygies but a stronger statement of Schreyer implies that they form a Gröbner basis. To build a bridge with Lazard's structure theorem, we observe that the minors can be split into two components

$$g_i = \underbrace{(\prod_{j=1}^{i-1} -x^{d_j} - n_{j,j})}_{M_i} \underbrace{\det\left(\begin{bmatrix} y & -x^{D_{i+1}} & & \\ & \ddots & & \\ & & \ddots & -x^{D_{n_0}} \\ & & & y \end{bmatrix} + \begin{bmatrix} -n_{i+1,i} & -n_{i+1,i+1} & \\ \vdots & \ddots & \\ -n_{n_0+1,i} & \cdots & -n_{n_0+1,n_0+1} \end{bmatrix}\right)}_{G_i}$$

where $M_i = \prod_{j=0}^{i} D_j$, $i \in [0,s]$, for some $D_i \in \mathbb{K}[x]$ carries the the greatest $\mathbb{K}[x]$ divisor of g_i. The G_i part describes the y component with the membership condition in Lazard's structure theorem. Note that $P = 1$ in the structure theorem under the hypothesis that $|V(I)| < \infty$.

The map is also defined under some restrictions. For example, if we restrict

$$\mathcal{C}_0(\boldsymbol{E}) = C(\boldsymbol{E}) \cap \{I \mid V(I) = (0,0)\}$$

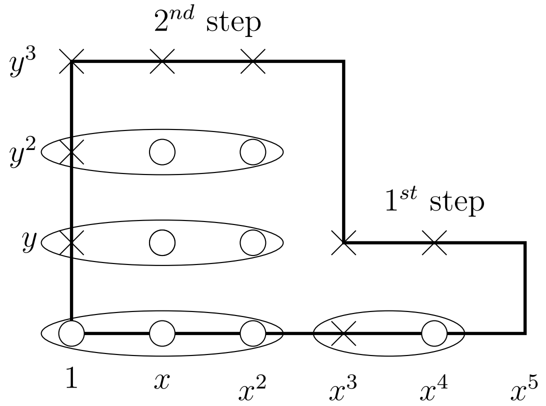
the parametric generators, the $g_i's$, are like above but with the additional restrictions that $n_{i,i} = 0$ for all i and $n_{i,j}(0) = 0$ if $i > j + w_i$, where w_i the *height of the stair* at y^i if any. We say that $\mathcal{C}_0(\boldsymbol{E})$ is a punctual Gröbner cell. In the case of this restriction number of free parameters in Λ is less than $\dim_{\mathbb{K}} \mathbb{K}[x,y]/\langle \boldsymbol{E} \rangle$, in other words, the dimension of the moduli space of $C_0(\boldsymbol{E})$ is at most the multiplicity at the origin of all ideals $I \in C_0(\boldsymbol{E})$.

❧ Example 2.1.26: Two Stairs Border Basis from [40]

Let

$$\boldsymbol{E} = (y^3, x^3y, x^5),$$

the elements in the canonical Hilbert-Burch matrix of $\mathcal{C}_0(\boldsymbol{E})$ can be described as



$$\boldsymbol{H} = \begin{pmatrix} 0 & 0 & 0 \\ -\Lambda_1 x - \Lambda_2 x^2 & 0 & 0 \\ -\Lambda_3 x - \Lambda_4 x^2 & 0 & 0 \\ -\Lambda_5 - \Lambda_6 x - \Lambda_7 x^2 & 0 & -\Lambda_8 x \end{pmatrix} + \begin{pmatrix} -x^3 & 0 & 0 \\ y & -1 & 0 \\ 0 & y & -x^2 \\ 0 & 0 & y \end{pmatrix}.$$

Using the result of [40], a paramatric Gröbner basis of $\mathcal{C}_0(\boldsymbol{E})$ is given by the 3-minors of $\boldsymbol{H}$:

$$\begin{aligned}&(y^3 - \Lambda_2 x^2 y^2 + (-\Lambda_1 - \Lambda_8) x y^2 + (\Lambda_1 d_8 - \Lambda_4) x^2 y - \Lambda_3 x y + \\ &(\Lambda_2 \Lambda_8^2 - \Lambda_7) x^4 + (\Lambda_4 \Lambda_8 - \Lambda_6) x^3 + (\Lambda_3 \Lambda_8 - \Lambda_5) x^2, \\ &x^3 y - \Lambda_8 x^4, \\ &x^5)\end{aligned}$$

In particular, the ponctual Gröbner cells $\mathcal{C}_0(\boldsymbol{E})$ is the affine variety $\mathbb{A}^8(\mathbb{K})$. Each ideal in $\mathcal{C}_0(\boldsymbol{E})$ corresponds to a unique point $(p_1, \ldots, p_8) \in \mathbb{A}^8(\mathbb{K})$; when replacing Λ_i by p_i in the minors, we obtain a Gröbner basis of of I.

2.2 State of the Art

Algebraic methods to describe the vanishing of I are numerous and vary depending on the number of variables involved and the ambient field. They notably include linear algebra, factorisation, triangular representations, homotopy, Gröbner basis of the ideal I (non-exhaustive list). Here, we emphasise that we not only want to reduce the problem of describing $\operatorname{Spec}\mathbb{K}[x,y]/I$, *i.e.* the intersections and their respective local structure, but we also want to get the best complexity to do so. Complexity is generally calculated in terms of the number of operations in the base ring, either in an abstract sense or in a smaller metric considering the elements' size. When considering the size, the complexity is more precise and is oft qualified as *bit-wise*[8].

Before jumping to the heart of the matter, let us review some basic notions of arithmetic complexity that will be used throughout the document.

2.2.1 Folklore of arithmetic complexity

2.2.1.0.1 Abstract operations on a ring Let $\mathbb{A} = (\mathbb{A}, +, \times, 1, 0)$ be a ring, then the arithmetic complexity in terms of the number of operations in $\mathbb{A}$ counts the operations such as $+$,$-$,$\times$, and $\div$ when feasible.

Let $R = \mathbb{A}[x_1, \dots, x_n]$ and $f_1, f_2 \in R$ then f_1+f_2 or f_1-f_2 take $\tilde{O}(\max\{\deg(f_i)^n\})$ operations in $\mathbb{A}$.

Multiplying in R satisfies a super-linearity condition for the number of operations in the base ring; notably, it can be done via Fast Fourier Transform (FFT) in [20, §2.4, Algorithm 2.2]. Euclidean algorithm, *i.e.* division with remainder, with polynomials in $\mathbb{K}[x]$ of degree less than d, for $\mathbb{K}$ a field, can be done via the fast extended division algorithm [73, §11.1, Corollary 11.6] or Newton iteration in $\tilde{O}(d)$ operations in $\mathbb{K}$. In the Newton approach, the quotient can be found by first defining an inverse of a divisor D, $D\xi_n \equiv 1 \mod x^{2n}$ (see Example 2.1.24) and then finding the quotient of the division of f by D as $Q = f\xi_n \mod x^{2n}$. Since the main ring has no zero divisors, the degree of the quotient, Q is upper bounded by $\deg(f) - \deg(D)$; so it is easy to choose when to stop the Newton iterations. The remainder can be found from f, Q and D. Inversion in $\mathbb{A} = k\left[t\right]/\langle f^i\rangle$, where f is irreducible in $k\left[t\right]$ with extended Euclidean algorithm in $\tilde{O}(i \deg f)$ operations in k [73, §11].

Let $\mathbb{A}^{n\times m}$ be the set of $n\times m$ matrices over the ring $\mathbb{A}$, then linear operations such as $+, -$ require $O(n \times m)$ operations in $\mathbb{A}$. If $\mathbb{A}$ is a field and $m = n$, multiplicating and, when possible: *finding the determinant; diagonalizing; finding the characteristic polynomial; and inversion*, can all be done in $O(n^\omega)$ operations in $\mathbb{A}$ where

[8] especially when $\mathbb{A} = \mathbb{Z}$

$2 \leq \omega < 3$ [20, §8.1]. The best-known value for ω to this day is $\omega = 2.37188\dots$ [57]. It is desirable, although it has never been proven, that ω could be 2. It is sometimes feasible to get better bounds for some matrices operations when the matrices are structured, *e.g.* Toeplitz, Hankel, companion, Vandermonde matrices can be inverted and multiplied with a vector in $O(M(n)\log n)$ operations, where $O(M(n))$ is the complexity of × in $\mathbb{A}[x]$ between polynomials of degree at most n [20, §33, Excercise 33.12].

2.2.1.0.2 Binary operations over finite rings When working over finite rings, *e.g.* $\mathbb{A} = \mathbb{Z}/p^i\mathbb{Z}$, $\mathbb{A} = k[t_1, \dots, t_n]/\mathfrak{m}^i$ for k a finite field and $\mathfrak{m} \subseteq k[t_1, \dots, t_n]$, complexity analysis sometimes take into account the cost of each operation in the base ring in term of operations in a smaller metric, *e.g.* binary operations.

❧ **Example 2.2.1: $\mathbb{A} = \mathbb{Z}/p^i\mathbb{Z}$**

When $\mathbb{A} = \mathbb{Z}/p^i\mathbb{Z}$ operations $+, -, \times$ and the Euclidean algorithm (division and modulo), when feasible, can be done in $O\tilde{}(i \log p)$ binary operations.

When we consider the cost of the fundamental operations in the ring, we often say that this is with regard to the size of the representation. The asymptotic complexity analysis is established on the number of operations in the ring weighted with the cost of fundamental operations in this ring

$$(\text{\# of abstract operations}) \times (\text{cost of one fundamental operation in the finite ring}).$$

In the analysis, we use the $O\tilde{}()$ to mark the omission of logarithmic factors.

2.2.2 History

Solving polynomial systems has acquired more than one meaning over time. It has a purely numerical sense, for which there is no hope of describing the local algebraic structure of the roots. Mostly algebraic algorithms, so-called *"exact"* solutions, are considered hereinafter. Following the observations of a Marinari, Möller *&* Mora in [122], we can further identify the **arithmetic sense** and the **scheme sense**, which the authors referred to as the *algebraic senses*.

2.2.2.1 The arithmetic sense

For univariate systems, factorisation gives a precise description of the solutions. Actually, in $\mathbb{K}[x_1, \ldots, x_n]$, most algebraic solutions reduce the problem of finding the locus to a factorisation problem. This is notably the case of the Shape Lemma, the subresultant/resultant (for bivariate systems), symbolic homotopy and Gröbner bases.

Lemma 2.2.1 (Shape Lemma)**.** *Under the below assumptions,* $\mathbf{H}_1$ *and* $\mathbf{H}_2$*, a finite set* $V(I) \subseteq \bar{\mathbb{K}}^2$ *in any algebraically closed field* $\bar{\mathbb{K}}$ *may be described by using a pair* (u, v) *of polynomials in* $\mathbb{K}[x]$ *such that* $V(I) = V(\langle u(x), y - v(x)\rangle)$ *and* u *is square free .*

$\mathbf{H}_1$**.** $\mathbb{K}$ *is a perfect field (i.e.: all irreducible polynomials in* $\mathbb{K}[x]$ *have pairwise distinct roots in* $\bar{\mathbb{K}}$*).*

$\mathbf{H}_2$**.** $V(I)$ *is in generic coordinates, in the sense that* $\forall\ \xi = (x, y), \xi' = (x', y') \in V(I) : x \neq x'$*.*

This property is better known as the Shape Lemma [76].

If (u, v) is a pair of the Shape Lemma for the zero-dimensional ideal I, then $\langle u(x), y - v(x)\rangle$ is a Gröbner basis of $I(V(I))$. Furthermore, for all $p \in V(\langle u(x), y - v(x)\rangle)$, p is smooth.

Some remarks on the hypotheses:

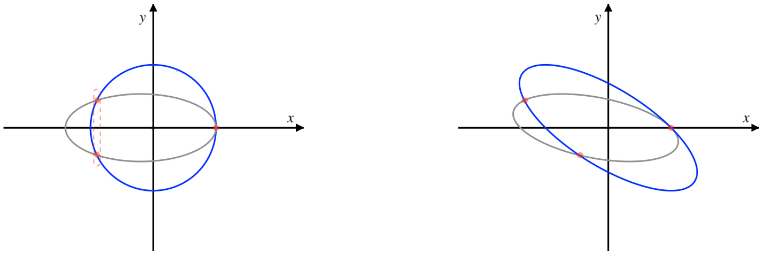
Remark 2.2.1. All fields of characteristic zero and all finite fields are perfect. Thus $\mathbf{H}_1$ holds for finite fields, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

Remark 2.2.2. In the case where the coordinates are non-generic, we may use the concept of *equiprojectable decomposition*, *i.e.*, grouping the solutions by the number of solutions that share the same abscissa [111]. In contrast, most recent algorithms prefer to put their system in generic coordinates with a reversible and inexpensive transformation such as ϕ defined below to find only a pair (u, v).

$$\phi : x \mapsto x + ty \quad \text{with } t \in \mathbb{Z}$$

❧ Example 2.2.2

Let $V(I) = V(ay^2 + (x+1)^2 - (a+1)^2) \cap V(y^2 + x^2 - a^2)$ with $a \in \mathbb{Z}$, an ellipse that intersects a circle. Then we have two families: one with one solution per abscissa and another with two solutions per abscissa. However, the variety is in generic coordinates when considering its image under the map $\phi : x \mapsto x + y$.

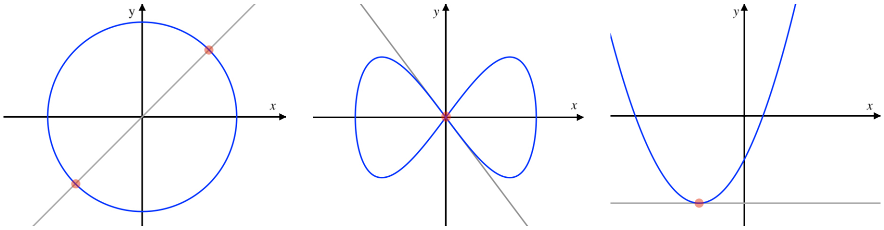


$V(\langle 5y^2 + (x+1)^2 - 6^2, y^2 + x^2 - 5^2\rangle)$ $V(\langle 5y^2 + (x+y+1)^2 - 6^2, y^2 + (x+y)^2 - 5^2\rangle)$

Figure 2.2: Example of change to generic coordinates.

❧ Example 2.2.3: Shape Lemma

Revisiting our intersections of multiplicity 1 and 2 from Example 1.1.1, the positions of the element of $V(\langle x-y\rangle) \cap (\langle x^2+y^2-1\rangle)$, $V(\langle y+x\rangle) \cap V(\langle y^2+x^4-x^2\rangle)$ or $V(\langle y+2\rangle) \cap V(\langle y-x^2-2x+1\rangle)$ can be characterized by the pairs $(2x^2-1, y-x)$, (x, y) and $(x+1, y+2)$ respectively,



$V(\langle x - y, x^2 + y^2 - 1)\rangle)$ $V(\langle y + x, y^2 + x^4 - x^2\rangle)$ $V(\langle y + 2, y - x^2 - 2x + 1\rangle)$

$\{(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\}$ $\{(0,0)\}$ $\{(-1,-2)\}$

$u(x) = 2x^2 - 1, v(x) = -x$ $u(x) = x, v(x) = 0$ $u(x) = x + 1, v(x) = 2$

Figure 2.3: Shape Lemma for Example 1.1.1.

where the roots of $u(x)$ give the abscissae of the points in $V(I)$ and $v(x)$ evaluated to the roots gives the ordinates. In the case of the intersection between 2 plane curves $V(f_1) \cap V(f_2)$ like the varieties aforementioned, if the degrees of the f_i's are at most d, then the degrees of $u(x)$ and $v(x)$ is *sharply* bounded by d^2 [124].

The Shape Lemma can be found via a limit construction approach. We first focus on the work done over the rational numbers or the function field of an affine line, *i.e.*, $k(t)$ for k a field, where the points in the variety are all smooth. In 1984, Trinks [153] gave a p-adic algorithm, for p a prime number, to find the multivariate equivalent of the Shape Lemma for polynomials over $\mathbb{Q}$ in n variables, *i.e.*, $(g_1, \ldots g_n)$ such that $V(I) = V(\langle x_1 - g_1(x_n), \ldots x_{n-1} - g_{n-1}(x_n), g_n(x_n)\rangle)$. Later, Giusti *& al.* [82, 79, 81, 83] rediscovered and filled details of the lifting. In 2003, Schost [142] presented a probabilistic Newton iterator to find a triangular decomposition with a complexity subquartic with respect to the size of the decomposition. In 2013, in collaboration with Merahbi *&* Lebreton [111], they concluded the result in the bivariate setting for two generators. They give an algorithm to find the equiprojectable decomposition, an adaptation on Shape Lemma such that $\mathbf{H}_2$ is not required, for $f_1, f_2 \in \mathcal{S}[x, y]$, each of total degree at most d, with a probability of 1/2 using

- $\tilde{O}(d^{3.69}h + d^{4.69})$ operations in $\mathbb{K}$ if $\mathcal{S} = \{f \in k[t] \mid \deg f \leq h\}$ where $k[t]$ is the function field of the affine line $\mathbb{A}^1(k)$ for k a field [111, Theorem 1];
- $\tilde{O}(d^{3+\varepsilon}h + d^{4+\varepsilon})$ bit operations for any $\varepsilon > 0$ if $\mathcal{S} = \{z \in \mathbb{Z}, |z| < h\}$ [111, Theorem 2] .

All the above work exclusively on smooth points. Lecerf's deflation algorithm [112] reestablishes Newton iteration's quadratic convergence in the occurrence of multiple roots. In 2016, the idea was revisited for bivariate ideals by Mehrabi and Schost [124]. Their work is notably focused on modular operations to showcase a complexity of $d^{2\varepsilon}\tilde{O}(d^2 + dh + dP + P^2)$ bit operations, for any $\varepsilon > 0$, to find the Shape Lemma of $f, g \in \mathcal{S}[x, y]_{\leq d}$ with a probability of $1 - \frac{1}{2}^P$ for $P \in \mathbb{Z}$ and $\mathcal{S} = \{z \in \mathbb{Z} \mid |z| < h\}$ [124].

The Shape Lemma and variant accurately describe $V(I)$ set-wise. However, it conceals geometric information at the intersection.

❧ **Example 2.2.4**

For example, $V(\langle x^3, y\rangle)$ and $V(\langle x^2, xy, y^2\rangle)$ are both equivalent to $V(\langle x, y\rangle)$. Thus evidently, by Example 2.1.7, the pair does not suffice to verify local isomorphisms.

As underlined in Chapter 4, there exists a refinement to the Shape Lemma; by removing the restriction on u, one could enrich elements of $V(I)$ with the multiplicity of the intersections [140]. However, as the previous Example 2.2.4 and Example 2.1.7 illustrated, it does not suffice to preserve the local structure of a point.

2.2.2.2 The scheme sense

For a fair comparison, we now focus on algorithms that lead toward a scheme description. Marinari, Möller & Mora [122] presented a detailed study of possible representations (*e.g.* Border bases, inverse systems, Gröbner bases) and they discussed how to pass from one representation to a dual. However, they did not provide an algorithm to find a Gröbner basis. The inverse system of an isolated root, *i.e.* a representation based on the partial derivative, can be found with Macaulay's dialytic method [119]. Mourrain [126] also described an approach to find an inverse system of an isolated root and, with Mantzaflaris & Szanto [121], they described a Newton method over $\mathbb{C}$ under regularity conditions. The system involved is overdetermined; therefore, they introduced a perturbation. Haustein, Mourrain & Szanto [88] presented an algorithm to find border bases, *i.e.* multiplication matrices for a basis of the quotient ring of an ideal $\langle \mathcal{F} \rangle \subseteq \mathbb{K}[x_1, \ldots, x_n]$ using $|\mathcal{F}|\delta + n(n-1)(\delta-1)(\delta-2)/4$ equations in $n + n\delta(\delta-1)/2$ variables where $\delta = \dim_{\mathbb{K}} \mathbb{K}[x_1, \ldots, x_n]/\langle \mathcal{F} \rangle$. Their technique can be adapted to a $\mathfrak{m}$-adic scheme.

We opted for the Gröbner bases (complete ideal or primary components) with a parametrisation of Gröbner cell as it leads to a simpler algorithm with fewer parameters than border bases while presenting additional applications.

2.2.2.2.1 Primary decompositions Based on his structure theorem (see Theorem 2.1.6), Lazard [109] presented an algorithm, complexity analysis excluded, to decompose bivariate ideal. His algorithm assumes that the associated primes are known. Algorithms also exist for ideals in $\mathbb{K}[x_1, \ldots, x_n]$ [53]. In general, the idea is to perform some preliminary steps, like finding a Gröbner basis and factorising polynomials, assuming we know how to factorise in the base ring.

2.2.2.2.2 Gröbner basis The BUCHBERGER'S ALGORITHM can be optimised notably by removing superfluous generators as we go, limiting the number of remainders evaluated, etc. As a result, the asymptotic complexity to find a Gröbner basis of a bivariate ideal $I = \langle \mathcal{F} \rangle \subseteq \mathbb{K}[x, y]$ can be bounded by $\frac{3}{2}(|\mathcal{F}| + 2(d+2)^2)^4$ operations in the field [34], where each $f \in \mathcal{F}$ has degree at most d. For complexity over finite rings, less is known. Over $\mathbb{Z}$, we can optimize based on the density of the generators to get a binary complexity that is *polynomial* in terms of the maximum between the arithmetic mean value degree and the maximum size of the generators in a dense representation [87].

In 1983, Ebert addressed the question of modular algorithms for Gröbner bases [59]. He proved that, over $\mathbb{Q}$, it is feasible to validate if a prime number is *lucky* in the restrictive context of binomials and monomials Gröbner basis. In 1988, still over $\mathbb{Q}$, Winkler gave the first p-adic algorithm to find a Gröbner basis [155]. Pauer, in 1992, gave a lifting and addressed the question of *lucky* prime numbers [135]. The idea was then simplified by Arnold in 2003 [6]. She offered a constructive iteration to obtain Gröbner bases of multivariate homogenous ideals. By extension, the iteration can easily be adapted to use a graded monomial ordering (see Example 2.1.10 and Example 3.1.4). However, she did not complete a complexity analysis of this algorithm. Each of Arnold's, and her predecessor's, iterations entail solving linear equations over $\mathbb{K}[x_1, \ldots, x_n]$[9]. The construction presented here requires primarily modular evaluations and a linear system in a base ring to be inverted.

State of the art, until then, did not address the question of whether a quadratic convergence was feasible for a limit construction of a Gröbner basis. In all the aforementioned algorithms, the nature of the iteration is not as simple as defining a nice variety on which a Newton iteration can be applied. Only linear convergence methods emerge from antecedent work.

[9] Thus each step solves something equivalent to a Gröbner basis, so to speak.

Chapter 3

Change of Basis for $\mathfrak{m}$-primary Ideals in One and Two Variables

Overview of this Chapter 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.

Published S. G. Hyun, S. Melczer, É. Schost, and C. St-Pierre. Change of basis for $\mathfrak{m}$-primary ideals in one and two variables. In *ISSAC'19*, pages 227–234. ACM Press, 2019 doi: 10.1145/3326229.332626

3.1 Introduction

In [91], 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 [146]. This saves logarithmic factors compared to the results in [91]; 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 [91] 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 [117, 35, 80, 5, 140], or triangular representations [156, 8]. 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 an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$.

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 [122], one may also give a basis of the dual of $\mathbb{F}[x_1, x_2]/I$, or a standard basis of I. In [122, 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 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 [128]. We take $\mathbb{F} = \mathbb{Q}$, $\mathfrak{m}$ to be 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 $I = \mathfrak{m}^2$ to be the $\mathfrak{m}$-primary ideal with generators

$$\begin{aligned} G_3 &= x_2^2 - 2x_1x_2 - 2x_2 + x_1^2 + 2x_1 + 1, \\ G_2 &= x_1^2x_2 + x_1x_2 + 2x_2 - x_1^3 - 2x_1^2 - 3x_1 - 2, \\ G_1 &= x_1^4 + 2x_1^3 + 5x_1^2 + 4x_1 + 4. \end{aligned}$$

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 to be the residue class of y_1 and $\alpha_2 = \alpha_1 + 1$.

The $\langle \alpha_1, \alpha_2 \rangle$-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 $\langle \xi_1 - \alpha_1, \xi_2 - \alpha_2 \rangle$, is the ideal with lexicographic Gröbner basis

$$\begin{aligned} H_3 &= \xi_2^2 - 2\xi_2\alpha_1 - 2\xi_2 + \alpha_1 - 1, \\ H_2 &= \xi_1\xi_2 - \xi_2\alpha_1 - \xi_1\alpha_1 - \xi_1 - 2, \\ H_1 &= \xi_1^2 - 2\xi_1\alpha_1 - \alpha_1 - 2. \end{aligned}$$

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. Note that with a strong regularity assumption and in a different setting, it has been shown in [90] that multiplication in $\mathbb{F}[x_1, x_2]/I$ can be done in quasi-linear time.

Although our results are still partial (we make assumptions and deal only with bivariate systems), we believe it is worthwhile to investigate these questions. In future work (Chapter 4 and 5), 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 [88] for instance, as well as the computation of GCDs modulo ideals such as I above. See, for instance, [47] for a discussion of the latter question.

3.2 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 $\mathbb{F}$.

3.2.1

We let $\mathsf{M} : \mathbb{N} \to \mathbb{N}$ be such that the 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 [73, Chapter 8]. Below, we will freely use all usual consequences of fast multiplication (on fast GCD, Newton's iteration, ...) and refer the reader to *e.g.* [73] 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).

3.2.1.1

We will use the *transposition principle* [36, 100], 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 [146, 23].

3.2.2

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 3.2.1, 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 [147] 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 3.4.4.

3.3 The univariate case revisited

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

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 [91] gives algorithms for both untangling and tangling, the latter calling the former recursively; the untangling algorithm runs in $O(\mathsf{M}(d\mu)\log(\mu))$ operations in $\mathbb{F}$, while the tangling algorithm takes $O(\mathsf{M}(d\mu)\log(\mu)^2 + \mathsf{M}(d)\log(d))$ operations. Using transposition techniques from [146], we prove the following.

Property 3.3.1. 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 $\mathcal{C} = (1, \ldots, \alpha^{d-1}, \xi, \ldots, \alpha^{d-1}\xi, \ldots \xi^{\mu-1}, \ldots, \alpha^{d-1}\xi^{\mu-1})$. As per 3.2.2, 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 $[\ell(\alpha^i \xi^j) \ \mid \ 0 \le i < d, \ 0 \le j < \mu] \in \mathbb{F}^{d \times \mu}$.

3.3.1 A faster tangling algorithm

This section shows that using the transpose of untangling allows us to deduce an algorithm for tangling; see [146, 52] for a similar use of transposition techniques. We start by describing useful subroutines.

3.3.1.1

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's iteration modulo the powers of T, at the cost of another $O(\mathsf{M}(d\mu))$.

3.3.1.2

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 $[H][f_0, \ldots, f_{d\mu-1}]^T = [B]$ with $H_{i,j} = L(x^{i+j})$ and $B_i = L'(x^i)$, $0 \le i < d\mu$. The system can be solved in $O(\mathsf{M}(d\mu)\log(d\mu))$ operations in $\mathbb{F}$ [31], but we will 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 [21, 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 3.3.1.1, 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 $O(\mathsf{M}(d\mu) + \mathsf{M}(d)\log(d))$.

3.3.1.3

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 $F, F', \dots, F^{(\mu-1)}$ at α, divided respectively by $0!, 1!, \dots, (\mu-1)!$.

Algorithm 3.3.1 $\pi_{\mathrm{rec}}(F, T, \mu)$

```
INPUT: $F \in \mathbb{F}[x]/\langle T^\mu\rangle$
OUTPUT: $[F(\alpha), \dots, F^{(\mu-1)}(\alpha)] \in \mathbb{K}^\mu$
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 3.3.2 $\pi(F, T, \mu)$

```
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 $\mathcal{T}(d,\mu)$ of π_{rec} satisfies $\mathcal{T}(d,\mu) \le \mathcal{T}(d,\mu/2) + O(\mathsf{M}(d\mu))$, 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 3.3.3 $\pi_{\mathrm{rec}}^{\perp}(\ell, T, \mu)$

```
INPUT: $\ell \in \mathbb{F}^{d\times\mu}$
  if $\mu = 1$ then return $\ell_0$ 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 3.3.4 $\pi^{\perp}(\ell, T, \mu)$

```
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 [23], 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 $i = \lambda, \dots, t-1$. It can be computed in linear time $O(t)$. Overall, as in [91], the runtime is $O(\mathsf{M}(d\mu)\log(\mu))$.

3.3.1.4

We can now give our algorithm for the tangling operator ${\pi_{T,\mu}}^{-1}$; it is inspired by a similar result due to Shoup [146].

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 3.2.2, 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 3.2.2, 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 3.3.1.2; 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 3.3.1.

3.3.2 Applications

3.3.2.1

For P in $\mathbb{F}[x]$ one can compute $x^D \bmod P$ using $O(\log(D))$ multiplications modulo P by repeated squaring. Applications include Fiduccia's algorithm for the computation of terms in linearly recurrent sequences [70] or of high powers of matrices [137, 77]. This algorithm takes $O(\mathsf{M}(n)\log(D))$ operations in $\mathbb{F}$, with $n := \deg(P)$. We assume without loss of generality that $D \geq n$.

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 $Q(x,0)$ [19].

First, assume that $P = T^\mu$ with T separable of degree d. Then we compute $x^D \bmod P$ 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 3.3.1, 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 $d \leq D$. 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 only for astronomical values of the parameters.

3.3.2.2

Assume $\mathbb{F} = \mathbb{Q}$. In [111], 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 [111] 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 [111] is the best to date.

The case of non-radical systems was discussed in [124]. 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 [111] 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 a radicality assumption.

3.4 The bivariate case

We now generalize the previous questions to the bivariate setting. 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.

3.4.1 Setup

3.4.1.1

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:

$\mathbf{H}_2$**.** $\mathbb{F}$ 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 $(0, 0)$.

3.4.1.2

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 3.4.1.1.

3.4.1.3

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 above).

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 $d_1 d_2 = d = \deg(\mathfrak{m})$. 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.

Property 3.4.1. 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 sub-sections are devoted to the proof of this proposition.

3.4.1.4

We define here a family of polynomials $G_1, \ldots, G_t$, and prove that they form a (non-reduced) Gröbner basis of I in 3.4.1.5.

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 $\beta_1 = \beta, \beta_2, \ldots, \beta_d$ be the roots of F in $\mathbb{L}$. For all $i = 1, \ldots, d$, 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, $\alpha^{(1)}, \ldots, \alpha^{(d)}$ can be seen as all the roots of $\mathfrak{m}$, with $\alpha^{(1)} = \alpha$.

For $i = 1, \ldots, d$, 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 $1, \ldots, d$, 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 $(H_{i,1}, \ldots, H_{i,t})$, 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 $\{\alpha^{(1)}, \ldots, \alpha^{(d)}\}$ of $\mathfrak{m}$ according to their first coordinate, into d_1 classes $C_1, \ldots, C_{d_1}$ 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 $\zeta_1, \ldots, \zeta_{d_1}$ 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 [109] implies that for $j = 1, \ldots, t$, $H_{i,j} = (\xi_1 - \zeta_\kappa)^{\mu_j} \tilde{H}_{i,j}$, 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 $(G_1, \ldots, G_t)$, 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.

3.4.1.5

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

Take such an index i, and assume that $\alpha^{(i)}$ is in C_κ, for some $\kappa \leq d_1$ (so the first coordinate of $\alpha^{(i)}$ is ζ_κ). Take D large enough, so that $D \geq \mu_1$ and $(\xi_1 - \zeta_\kappa)^D$ belongs to Q_i; hence Q_i is also the primary component of the ideal $\langle G_1, \ldots, G_t, (\xi_1 - \zeta_\kappa)^D \rangle$ 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 $2 \leq j \leq t$. 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 $(\xi_1 - \zeta_\kappa)^{\mu_1}, (\xi_1 - \zeta_\kappa)^{\mu_2} \tilde{G}_{\kappa,2}, \ldots, \tilde{G}_{\kappa,t}$. 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 $\iota \neq i$, $\tilde{H}_{\iota,j}$ does not vanish at $\alpha^{(i)}$ [109, Theorem 2.(i)], so it is invertible locally at $\alpha^{(i)}$. It follows that the primary component of G at $\alpha^{(i)}$ is generated by $(\xi_1 - \zeta_\kappa)^{\mu_1}, (\xi_1 - \zeta_\kappa)^{\mu_2} \tilde{H}_{i,2}, \ldots, \tilde{H}_{i,t}$, that is, $H_{i,1}, \ldots, H_{i,t}$. This is precisely the ideal K_i.

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

The polynomials $G_1, \ldots, G_t$ are a Gröbner basis, *minimal*, as can be seen from their leading terms, but not reduced; we let $R_1, \ldots, R_t$ 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.

3.4.1.6

As a corollary, the following proposition and its proof extend [91, 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 3.4.1. *Assume $\mathfrak{m}$ is a maximal ideal in $\mathbb{F}[x_1, x_2]$ and I is an $\mathfrak{m}$-primary zero-dimensional ideal in $\mathbb{F}[x_1, x_2]$, with $\mathbb{F}$ perfect of characteristic at least $\deg(I)$.*

Let $\tilde{\mathfrak{m}}$ be the image of $\mathfrak{m}$ through the isomorphism $\mathbb{F}[x_1, x_2] \simeq \mathbb{F}[y_1, y_2]$, let α_1, α_2 be the residue classes of y_1, y_2 in $\mathbb{K} := \mathbb{F}[y_1, y_2]/\tilde{\mathfrak{m}}$ and let J be the primary component of $I \cdot \mathbb{K}[\xi_1, \xi_2]$ at (α_1, α_2). Finally, let J' be the image of J through $(\xi_1, \xi_2) \mapsto$

$(\xi_1+\alpha_1, \xi_2+\alpha_2)$. *Then, there exists an* $\mathbb{F}$*-algebra isomorphism*

$$\pi_{\mathfrak{m},J'} : \mathbb{F}[x_1,x_2]/I \to \mathbb{K}[\xi_1,\xi_2]/J' \tag{3.4.0.1}$$

given by $(x_1,x_2) \mapsto (\xi_1+\alpha_1, \xi_2+\alpha_2)$

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. □

3.4.2 Untangling for monomial ideals

3.4.2.1

In this section, we give an algorithm for the mapping $\pi_{\mathfrak{m},J'}$ of Proposition 3.4.1 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 3.4.1.3 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 3.4.1 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.

Property 3.4.2. Under **H**$_2$ and **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 3.4.2.2 and 3.4.2.3 respectively. The last statement readily follows, since $n = d\mu$ (Proposition 3.4.1).

3.4.2.2

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 [2]). Using Kronecker substitution [73, 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 [74] 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 3.4.1, 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 $n = \deg(I)$.

3.4.2.3

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 3.4.1, 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 $O(\mu_1\mathsf{M}(d_1d_2\nu_t)\log(\nu_t))$ 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).

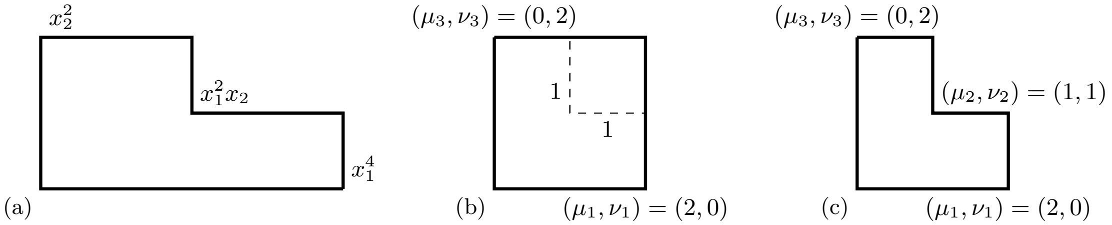


Figure 3.1: Monomials through untangling when μ is small.

3.4.3 Recursive tangling for monomial ideals

The ideas used to perform univariate tangling, that is, to 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 the 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 3.4.1.

Property 3.4.3. 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 [91], 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 $\mu_1 = 1$.

3.4.3.1

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 $F(\alpha_1, x_2) := \pi^{-1}_{T_2,\nu_2}(G) \in \mathbb{K}'[x_2]/\langle T_2^{\mu_2} \rangle$, working 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.

3.4.3.2

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 closely mirror Algorithm 9 in [91]. For the cost analysis, we let $S(\mathfrak{m}, J')$ be the cost of applying $\pi_{\mathfrak{m},J'}$ (see Proposition 3.4.2) 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 $H := (G - \pi_{\mathfrak{m},J'}(\bar{F})) \operatorname{div} {\xi_1}^{\bar{\mu}}$, 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 our 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).

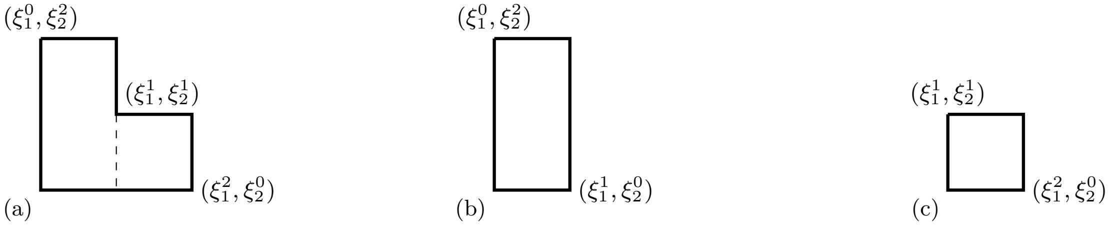


Figure 3.2: Monomials through recursive untanting.

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, $\pi_{\mathfrak{m},J'}(\bar{F}) = G \bmod {\xi_1}^{\bar{\mu}} + {\xi_1}^{\bar{\mu}}(\pi_{\mathfrak{m},J'}(\bar{F}) \operatorname{div} {\xi_1}^{\bar{\mu}})$. Using the definition of H, this is also $G \bmod {\xi_1}^{\bar{\mu}} + {\xi_1}^{\bar{\mu}}(G \operatorname{div} {\xi_1}^{\bar{\mu}} - H)$, 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 $\mathcal{T}(\mathfrak{m}, J')$ satisfies the recurrence relation

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

Using 3.4.3.1 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 $\mathcal{T}(\mathfrak{m}, J')$ is in $O(S(\mathfrak{m}, J')\log(n) + \mathsf{M}(n)\log(n)^2)$.

3.4.4 Tangling for monomial ideals using duality

We finally present a bivariate analogue of the algorithm introduced in Section 3.3. 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 [128], we now use several linear forms. Thus, let $\ell_1, \ldots, \ell_\gamma$ 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

$$L_1 := \pi^\perp_{\mathfrak{m},J'}(\ell_1), \dots, L_\gamma := \pi^\perp_{\mathfrak{m},J'}(\ell_\gamma)$$
$$L'_1 := \pi^\perp_{\mathfrak{m},J'}(\ell'_1), \dots, L'_\gamma := \pi^\perp_{\mathfrak{m},J'}(\ell'_\gamma)$$

in $(\mathbb{F}[x_1, x_2]/I)^*$. As in the one variable case, 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.* [61, 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 3.4.3. 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* [16]. It can be solved using structured linear algebra techniques [16] (Here, we have several such systems to solve at once; this can be dealt with as in [39]). As in [16], using the results from [21] 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 $\omega \le 2.38$ [43, 110]). 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.

3.4.5 An 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 3.4.3, 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.

The below section presents the original appendix of the publication.

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 this appendix, we write $I = \langle x_1^{\mu_1} x_2^{\nu_1}, x_1^{\mu_2} x_2^{\nu_2}, \ldots, x_1^{\mu_t} x_2^{\nu_t} \rangle$, with μ_i's decreasing, ν_i's increasing and $\nu_1 = \mu_t = 0$.

Property 3.4.4. 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 $i = 1, \ldots, t-1$; 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 $i_2 \leq t$ 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 $\mu_i \geq \mu_{i_r}/2$, 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, $\delta \geq 1/2\delta'$, that is, $\delta' \leq 2\delta$. To estimate the number s, the inequalities $\mu_{i_{r+1}} < \mu_{i_r}/2$ for all $r \leq s$ imply that $\mu_{i_{s-1}} < \mu_1/2^s$. We deduce that $2^s \leq \mu_1/\mu_{i_{s-1}} \leq \mu_1$ (since $\mu_{i_{s-1}} \geq 1$), 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 3.4.1. *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's 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's 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 $\mu_i \geq \bar{\mu}$ 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's iteration allows us to compute $G := 1/F \bmod I$ as $G = 2\bar{G} - \bar{G}^2F \bmod I$. 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 $O(\mathsf{M}(\delta)\log(\delta)^2)$.

Chapter 4

Newton iteration for lexicographic Gröbner bases in two variables

Overview of this Chapter We present an $\mathfrak{m}$-adic Newton iteration with quadratic convergence for lexicographic Gröbner basis of zero-dimensional ideals in two variables. We rely on a structural result about the syzygies in such a basis due to Conca *&* Valla, that allowed them to explicitly describe these Gröbner bases by affine parameters; our Newton iteration works directly with these parameters.

Preprint É. Schost and C. St-Pierre. Newton iteration for lexicographic gröbner bases in two variables, 2023. preprint on arXiv:2302.03766

4.1 Introduction

Solving bivariate polynomial equations plays an important role in algorithms for computational topology or computer graphics. As a result, there exists a large body of work dedicated to this question, using symbolic, numeric or mixed symbolic-numeric techniques [84, 67, 56, 4, 141, 14, 66, 27, 111, 25, 102, 124, 103, 26, 54, 48].

In many instances, these algorithms find a set-theoretic description of the solutions of a given system $f_1, \ldots, f_t$ in $\mathbb{K}[x, y]$ (here, $\mathbb{K}$ is a field). This can notably be done through the *shape lemma*: in generic coordinates, the output is a pair of polynomials u, v in $\mathbb{K}[x]$, with u squarefree, such that $V(\langle f_1, \ldots, f_t\rangle)$ is described by

$u(x) = 0$ and $y = v(x)/u'(x)$ (this rational form for y allows for a sharp control of the bit-size of v, if $\mathbb{K} = \mathbb{Q}$). One could slightly enrich this set-theoretic description by lifting the requirement that u be squarefree, and instead assign to a root ξ of u, corresponding to a point (ξ, ν), the multiplicity of $J = \langle f_1, \ldots, f_t \rangle$ at (ξ, ν) (adapting the definition of v accordingly). This is what is done in Rouillier's Rational Univariate Parametrization [140], but this still only gives partial information: for instance it is not sufficient to detect local isomorphisms.

In order to describe the solutions of J, but also the local structure of J at these zeros, it is natural to turn to Gröbner bases. This is what we do in this paper, our focus being an $\mathfrak{m}$-adic approximation procedure.

4.1.1 Our problem and our main result Let us assume that our base field $\mathbb{K}$ is the field of fractions of a domain $\mathbb{A}$, and take $f_1, \ldots, f_t$ in $\mathbb{A}[x, y]$.

Consider further the ideal $J = \langle f_1, \ldots, f_t \rangle$ in $\mathbb{K}[x, y]$. We are interested in finding a Gröbner basis of J itself, or possibly of some specific primary components of it. We will thus let I be an ideal in $\mathbb{K}[x, y]$, which we assume to be the intersection of some of the zero-dimensional primary components of J: typical cases of interest are $I = J$, if it has dimension zero, or I being the $\langle x, y \rangle$-primary component of J, if the origin is isolated in $V(J)$.

We let $\mathcal{G} = (g_0, \ldots, g_s)$ be the minimal, reduced Gröbner basis of I for the lexicographic order induced by $y \succ x$; this is the object we are interested in.

❧ **Example 4.1.1**

Let $\mathbb{A} = \mathbb{Z}$, and thus $\mathbb{K} = \mathbb{Q}$, $t = 2$ and input polynomials

$$f_1 = -12xy^5 - 20x^2y^4 - 14y^4 - 7x^3y^3 - 3x^2y^2 + 13x^3y - 17xy + 34x^2$$
$$f_2 = -x^2y^4 - 19x^3y^3 + 18xy^3 + 22x^3y^2 + 2x^2y^2 - 10x^2y.$$

We let I be the $\langle x, y \rangle$-primary component of $\langle f_1, f_2 \rangle$; its Gröbner basis $\mathcal{G}$ is

$$\left| \begin{array}{l} y^4 + \frac{17}{14}xy - \frac{17}{7}x^2, \\ xy^3 - \frac{10}{9}x^3, \\ x^2y - 2x^3, \\ x^4. \end{array} \right. \tag{4.1.0.1}$$

Let now $\mathfrak{m}$ be a maximal ideal in $\mathbb{A}$, with residual field $\Bbbk = \mathbb{A}/\mathfrak{m}$. Starting from the reduction of $\mathcal{G}$ modulo $\mathfrak{m}$ (assuming it is well-defined), the goal of this paper is to show how to recover $\mathcal{G}$ modulo powers of $\mathfrak{m}$. The case $\mathbb{A} = \mathbb{Z}$ seen above is the fundamental kind of example; another important situation is the "parametric" case, with $\mathbb{A} = \Bbbk[t_1, \dots, t_m]$ and $\mathfrak{m}$ a maximal ideal of the form $\langle t_1 - \tau_1, \dots, t_m - \tau_m \rangle$.

Let $\mathbb{A}_\mathfrak{m}$ ($\mathbb{A}_\mathfrak{m} \subseteq \mathbb{K}$) be the localization of $\mathbb{A}$ at $\mathfrak{m}$. For $K \geq 0$, there exists a well defined reduction operator $\mathbb{A}_\mathfrak{m} \to \mathbb{A}/\mathfrak{m}^K$, which we write $c \mapsto c \text{ rem } \mathfrak{m}^K$; we extend it coefficient-wise to a reduction mapping $\mathbb{A}_\mathfrak{m}[x, y] \to \mathbb{A}/\mathfrak{m}^K[x, y]$, and further to vectors of polynomials.

❧ **Definition 26.** *We say that* $\mathfrak{m}$ *is* **good** *with respect to* $f_1, \dots, f_t$ *and* $\mathcal{G}$ *if the following holds:*

- *all elements in* $\mathcal{G}$ *are in* $\mathbb{A}_\mathfrak{m}[x, y]$*,*
- *the ideal generated by* $\mathcal{G} \text{ rem } \mathfrak{m}$ *in* $\Bbbk[x, y]$ *is the intersection of some of the primary components of the ideal* $\langle f_1 \text{ rem } \mathfrak{m}, \dots, f_t \text{ rem } \mathfrak{m} \rangle$*.*

In particular, if $\mathfrak{m}$ is good, we will write $\mathcal{G}_\mathfrak{m}$ for the reduction $\mathcal{G} \text{ rem } \mathfrak{m}$. These are polynomials in $\Bbbk[x, y]$, and they still form a minimal, reduced Gröbner basis for the lexicographic order $y \succ x$.

❧ **Example 4.1.2**

In Example 4.1.1, $\mathfrak{m} = \langle 11 \rangle$ is good with respect to $f_1, \dots, f_t$ and $\mathcal{G}_\mathfrak{m}$ is

$$\left| \begin{array}{l} y^4 + 2xy + 7x^2, \\ xy^3 + 5x^3, \\ x^2y + 9x^3, \\ x^4. \end{array} \right.$$

If $\mathbb{A} = \mathbb{Z}$, there are finitely many primes p for which this is not the case. In the case $\mathbb{A} = \Bbbk[t_1, \dots, t_m]$, all maximal ideals of the form $\langle t_1 - \tau_1, \dots, t_m - \tau_m \rangle$ are good, except for those $(\tau_1, \dots, \tau_m)$ lying on a certain hypersurface in $\Bbbk^m$ (a quantitative analysis of the number of bad maximal ideals will be the subject of future work).

Our main result is an efficient lifting procedure based on Newton iteration to compute $\mathcal{G} \text{ rem } \mathfrak{m}^K$, given $f_1, \dots, f_t$, $\mathcal{G}_\mathfrak{m}$ and K. Lifting methods are widely used in computer algebra, for instance to solve linear systems or compute polynomial GCDs, and serve two purposes. First, while the arithmetic cost of solving the problem at hand (here, computing the Gröbner basis of I) may be high, our result will show that lifting an approximate solution modulo powers of $\mathfrak{m}$ is a relatively simple problem. Second, these techniques are usually used in cases where elements in $\mathbb{A}$, and $\mathbb{K}$, have a natural notion of "size" (such as the height when $\mathbb{A} = \mathbb{Z}$, or degree when $\mathbb{A} = \Bbbk[t_1, \dots, t_m]$). Then, direct computations in $\mathbb{K}$ often induce a significant "intermediate expression swell", where polynomials computed throughout the algorithm may have larger coefficients than the final output; $\mathfrak{m}$-adic approximation schemes avoid this issue.

Our algorithm features the quadratic convergence typical of Newton iteration, in the sense that it computes $\mathcal{G}$ rem $\mathfrak{m}^2, \mathcal{G}$ rem $\mathfrak{m}^4, \dots$; hence, without loss of generality, we assume that $K = 2^\kappa$ is a power of two. The cost of the algorithm is expressed in terms of two kinds of quantities:

- number of operations in the rings $\mathbb{A}/\mathfrak{m}^{2^i}$ (for which we discuss our computational model in more detail at the end of the introduction)
- the cost of reducing the coefficients of the polynomials f_j modulo $\mathfrak{m}^{2^i}$: we will assume that for $i \geq 0$, each such coefficient can be reduced modulo $\mathfrak{m}^{2^i}$ in time T_{2^i} (for $\mathbb{A} = \mathbb{Z}$, this time would depend on the bit-size of these coefficients; over $\mathbb{A} = \Bbbk[t_1, \dots, t_m]$, it would depend on their degree, and the number m of parameters).

Throughout, the $O\tilde{}$ notation indicates that we omit polylogarithmic factors, and ω is a feasible exponent for linear algebra.

Theorem 4.1.1. *Let $f_1, \dots, f_t$ be of degree at most d in $\mathbb{A}[x, y]$, with $\mathbb{A}$ a domain, that generate an ideal J in $\mathbb{K}[x, y]$, with $\mathbb{K}$ the fraction field of $\mathbb{A}$. Let I be the intersection of some of the zero-dimensional primary components of J, with minimal, reduced Gröbner basis $\mathcal{G}$, for the lexicographic order induced by $y \succ x$.*

Let further $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$ be the initial terms of $\mathcal{G}$, and let $\delta = \dim_{\mathbb{K}} \mathbb{K}[x, y]/I$.

Let $\mathfrak{m} \subseteq \mathbb{A}$ be a good maximal ideal for $\mathcal{G}$. For K of the form $K = 2^k$, given $\mathcal{G}$ rem $\mathfrak{m}$, one can find $\mathcal{G}$ rem $\mathfrak{m}^K$ with the following cost:

- *$O\tilde{}(s^2 n_0 m_s + t\delta(d^2 + dm_s + s\delta + \delta^{\omega-1}))$ operations in $\mathbb{A}/\mathfrak{m}^{2^i}$, for $i = 1, \dots, k$;*
- *$td^2 T_{2^i}$ steps for coefficient reduction, for $i = 1, \dots, k$.*

Remark 4.1.1. When I is the $\langle x, y\rangle$-primary component of J, runtimes can be sharpened, giving

- $O\tilde{}(s^2 n_0 m_s + t\delta^2(m_s + \delta^{\omega-2}))$ operations in $\mathbb{A}/\mathfrak{m}^{2^i}$, for $i = 1, \dots, k$;

- $t\delta m_s T_{2^i}$ steps for coefficient reduction, for $i = 1, \ldots, k$.

Since $m_s \leq \delta$, these are in particular $O\tilde{}(s^2 n_0 m_s + t\delta^3) \subset O\tilde{}((s+t)\delta^3)$, resp. $t\delta^2 T_{2^i}$. For the latter, we also have the bound $td^2 T_{2^i}$ stated in the theorem, but here we prefer to express the cost in terms of the multiplicity δ only.

This paper focuses on those cases where the ideal I is not radical (that is, where some points $p \in V(I)$ are singular), with the intent of computing the local structure at such points. If the sole interest is to find $V(I)$, then our approach is unnecessarily complex: the algorithms in [111, 124] use Newton iteration to compute a set-theoretic description of the solutions in an efficient manner.

❧ **Example 4.1.3**

An extreme case has $t = 2$ and f_1, f_2 "generic" in the sense that they define a radical ideal in $\mathbb{K}[x, y]$ with d^2 solutions in general position. In this case, if we take $I = J$, we have $s = 1$, $m_s = \delta = d^2$ and $n_0 = 1$. Then, the complexity in the first item of the theorem becomes $O\tilde{}(d^5)$ operations modulo each $\mathfrak{m}^{2^i}$. This is to be compared with the sub-cubic cost $O\tilde{}(d^{(\omega+3)/2})$ reported in [111] for a similar task.

Clearly, for these generic situations, our algorithm does not compare favourably with the state of the art. For the situation in Example 4.1.3, some techniques from [111] could be put to use in our situation as well, but they would at best give a runtime of $O\tilde{}(d^{2+(\omega+3)/2})$ operations in $\mathbb{A}/\mathfrak{m}^{2^i}$, still leaving a quadratic overhead. This is due to the different ways these papers apply Newton iteration: in our case, we linearize the problem in dimension d^2 (or, in, general, δ), and thus work with matrices of such size, whereas [111] work with matrices of size 2 (albeit with polynomial entries).

The results of Theorem 4.1.1 are of interest in the presence of intersection with multiplicities, where approaches such as [111] do not apply. The algorithm in [124] does not solve our problem in such cases, as it does not compute a Gröbner basis of I, but of its radical.

Remark that to derive a complete algorithm from our result, further ingredients are needed: quantitative bounds on the number of bad ideals $\mathfrak{m}$ (if $\mathbb{A} = \mathbb{Z}$ or $\mathbb{A} = \Bbbk[t_1, \ldots, t_m]$, for instance), a cost analysis for computing the starting point $\mathcal{G}_\mathfrak{m}$ and bounds on a sufficient precision K that will allow us to recover $\mathcal{G}$ from its approximation $\mathcal{G}$ rem $\mathfrak{m}^K$. In order to avoid this paper growing to an excessive length, we will address these questions in a separate manuscript.

We now review previous work on bivariate systems and Newton iteration for Gröbner bases. As we will see, there is a marked difference between Newton iteration algorithms for "simple" solutions (where the Jacobian of the input equations has full rank) in generic position and those that can handle arbitrary situations.

4.1.2 Newton iteration for non-degenerate solutions Following an early discussion in [59], p-adic techniques for Gröbner bases were introduced by Trinks in the 1980's [153]. That article focuses on zero-dimensional *radical* ideals with generators in $\mathbb{Z}[x_1, \ldots, x_n]$, in shape lemma position, that is, with a Gröbner basis of the form $x_1 - G_1(x_n), \ldots, x_{n-1} - G_{n-1}(x_n), G_n(x_n)$, for the lexicographic order $x_1 \succ \cdots \succ x_n$. Under this assumption, given a "lucky" prime p, one can apply a symbolic form of Newton iteration to lift $(G_1, \ldots, G_n)$ rem p to $(G_1, \ldots, G_n)$ rem p^K, for an arbitrary $K \geq 0$. Similar techniques were used in the geometric resolution algorithm [82, 81, 79, 83]; the scope of this symbolic form of Newton iteration was then extended in [142] to *triangular sets*, which are here understood as those particular lexicographic Gröbner bases $(G_1, \ldots, G_n)$ with respective initial terms of the form $x_1^{e_1}, \ldots, x_n^{e_n}$, for some positive integers $e_1, \ldots, e_n$. In [111], these techniques were studied in detail for the case $n = 2$ that concerns us in this paper, with a focus on the complexity of the lifting process.

Computationally, these algorithms are rather straightforward: they mainly perform matrix multiplications in size n with entries that are polynomials with coefficients in $\mathbb{Z}/p^K\mathbb{Z}$ (or more generally $\mathbb{A}/\mathfrak{m}^K$). These methods also share their numerical counterpart's quadratic convergence (in one iteration, the precision doubles, from p^K to p^{2K}), but none of them can directly handle solutions with multiplicities.

4.1.3 Lifting algorithms for general inputs [155] introduced an algorithm that handles arbitrary inputs: given a Gröbner basis $\mathcal{G}$ for $f_1, \ldots, f_t$ reduced modulo a "lucky" prime p, it recovers the Gröbner basis of the same system modulo p^K, for any $K \geq 0$. No assumption is made on the dimension of $V(\langle f_1, \ldots, f_t \rangle)$ or the rank of the Jacobian matrix of the equations. The computations are more complex as the ones above, as they involve lifting not only the Gröbner basis $\mathcal{G}$ itself, but also all quotients in the division of $f_1, \ldots, f_t$, and of the S-polynomials of $\mathcal{G}$, by $\mathcal{G}$.

In follow-up work, [135] discussed the choice of lucky primes; for homogeneous inputs, or graded orderings, [6] gave an efficient criterion to stop lifting and simplified the lifting algorithm itself, using ideas of Pauer's (the S-polynomials are not needed anymore).

To our knowledge, the algorithms mentioned here only perform *linear* lifting, going from an approximation modulo p^K to precision p^{K+1}; whether quadratic convergence is possible is unclear to us. No cost analysis was made.

4.1.4 Deflation Ojika, Watanabe and Mitsui introduced the idea of deflation in a numerical context [132], to restore Newton iteration's quadratic convergence even for multiple roots. The core idea is to replace the system we are given by another set of equations, having multiplicity one at the root we are interested in, possibly introducing new variables. There are now many references discussing this approach, see for instance [158, 112, 115, 116, 136, 51, 120, 157].

We are in particular going to use an idea from [88]. In that reference, Hauenstein, Mourrain and Szanto designed a deflation operator for an n-variate system $f_1, \ldots, f_t$, that converges quadratically to an augmented root (ξ, ν), where ν is a vector that specifies the local structure at a point $\xi \in V(\langle f_1, \ldots, f_t \rangle)$, through the coefficients of multiplication matrices in the local algebra at ξ. If ξ is known, this gives in particular an operator with quadratic convergence to compute the structure constants.

4.1.5 Our contribution The lifting algorithm we propose is simpler than in [155, 6] (we do not need to consider the polynomial quotients in the division of $f_1, \ldots, f_t$ by $\mathcal{G}$), but so far specific to lexicographic orders in two variables.

The first step is to identify a family of free parameters that describe Gröbner bases with given initial terms (these Gröbner bases form a *Gröbner cell*). The coefficients that appear in the Gröbner basis do not form such a family, as there are nontrivial relations between them. However, for lexicographic orders in two variables, Conca and Valla explicitly constructed a one-to-one parametrization of a given Gröbner cell by an affine space [40], from a description of canonical generators of the syzygy module. Our Newton iteration computes the parameters corresponding to $\mathcal{G}_\mathfrak{m}$ and lifts them modulo $\mathfrak{m}^K$.

This is done by adapting the approach of [88]: the coefficients of the normal forms of $f_1, \ldots, f_t$ modulo the unknown Gröbner basis $\mathcal{G}$ are polynomials in the parameters of the Gröbner cell; we prove that they admit as a (not necessarily unique) solution the parameters corresponding to $\mathcal{G}$, and that their Jacobian matrix has full rank at this solution. We can then apply Newton iteration to these polynomials, using $\mathcal{G}_\mathfrak{m}$ to give us their solution modulo $\mathfrak{m}$ as a starting point.

Computationally, the core operation is simply reduction modulo a lexicographic Gröbner basis. While we have algorithms with quasi-linear cost for reduction modulo a single polynomial, or modulo two polynomials with respective initial terms y^n and x^m, we are not aware of specific results for arbitrary lexicographic bases. Another contribution of this paper is a reduction algorithm, where we use techniques developed by van der Hoeven and Larrieu [90] for certain weighted orderings, adapted to our purposes.

4.1.6 Leitfaden In Section 4.2, we discuss *initial segments* in $\mathbb{N}^2$; they allow us to describe polynomials reduced modulo a Gröbner basis. We give in particular an algorithm for multiplying two such polynomials.

In Section 4.3, we review known results on the structure of bivariate lexicographic Gröbner bases: Lazard's theorem [109], and Conca and Valla's description of Gröbner cells; Section 4.4 then presents our algorithm for reduction modulo a lexicographic Gröbner basis.

In Section 4.5 and Section 4.6, we give algorithms to compute the Gröbner basis corresponding to a set of parameters in the Gröbner cell, and conversely. Finally, we describe Newton iteration for the Gröbner cell parameters in Section 4.7, proving Theorem 4.1.1.

4.1.7 Computational model In the whole paper, the costs of algorithms are measured using numbers of operations in the base ring or base field.

We will first and foremost count $\mathbb{Z}$-*algebra operations*. For an algorithm with inputs and outputs in a (unital) ring $\mathbb{A}$, these are additions and multiplications involving the inputs, previously computed quantities, and constants taken from the image of the canonical mapping $\mathbb{Z} \to \mathbb{A}$ (e.g., integers if $\mathbb{A}$ has characteristic zero); they will be simply be called "$(+, \times)$ operations". If an algorithm performs only this kind of operations, its outputs are in the subring of $\mathbb{A}$ generated by its inputs.

Important examples are addition, multiplication and Euclidean division (by a monic divisor) in $\mathbb{A}[x]$; they can all be done using a softly linear number of $(+, \times)$ operations in $\mathbb{A}$, over any base ring $\mathbb{A}$. For background, see Chapters 8 and 9 in [73].

Other operations we will occasionally use are invertibility tests and inversions (to solve linear systems). Finally, if $\mathfrak{m}$ is an ideal in a ring $\mathbb{A}$, given a in $\mathbb{A}/\mathfrak{m}$, we assume that we can find A in $\mathbb{A}$ with $A \text{ rem } \mathfrak{m} = a$ using one operation in $\mathbb{A}$.

4.1.8 Notation The following notation is used throughout the paper. In the following items, $\mathbb{A}$ is an arbitrary ring.

- For $d \geq 1$, We let $\mathbb{A}[x]_{<d}$ be the free $\mathbb{A}$-module of all polynomials in $\mathbb{A}[x]$ of degree less than d.

- For f, g in $\mathbb{A}[x]$, with f monic, we define $f \text{ rem } g$ and $f \text{ div } g$ as respectively the remainder and quotient in the Euclidean division of f by g.

- For f in $\mathbb{A}[x,y]$, $\deg(f,x)$ and $\deg(f,y)$ respectively denote its partial degrees with respect to x and y.

- For f in $\mathbb{A}[x,y]$ and $i \geq 0$, the *polynomial coefficient* of y^i in f will refer to the coefficient f_i in the expression $f = \sum_{i=0}^{d} f_i y^i$, with $f_0, \ldots, f_d$ in $\mathbb{A}[x]$. In the pseudo-code, we write $\textsc{PolynomialCoefficient}(f, y^i) \in \mathbb{A}[x]$ for this polynomial coefficient.

- If $f \in \mathbb{A}[x,y]$ has degree d in y, we say that f is *monic in y* if the polynomial coefficient of y^d is 1 (this definition and the previous one carry over to coefficients with respect to x instead, but we will not need this).

- If T is a subset of $\mathbb{N}^2$, we write $\mathbb{A}[x,y]_{\mathsf{T}}$ for the $\mathbb{A}$-module of polynomials *supported on* T, that is, all polynomials of the form $\sum_{(u,v)\in\mathsf{T}} a_{u,v} x^u y^v$, with only finitely many non-zero coefficients $a_{u,v}$.

We will not need to define Gröbner bases over rings. In particular, for reduction of bivariate polynomials, we only work over fields: if $\mathcal{G}$ is a Gröbner basis in $\mathbb{K}[x,y]$, where $\mathbb{K}$ is a field and $\mathbb{K}[x,y]$ is endowed with a monomial order, $f \operatorname{rem} \mathcal{G}$ denotes the remainder of f through reduction by $\mathcal{G}$.

4.2 Initial segments in $\mathbb{N}^2$

In this section, we first introduce terminology and basic constructions regarding subsets of $\mathbb{N}^2$ called initial segments. In the second part, we give algorithms to multiply polynomials supported on such initial segments.

4.2.1 Basic definitions

4.2.1.1 Initial segments We say that a set $\mathsf{T} \subset \mathbb{N}^2$ is an *initial segment* if for all (m,n) in T, any pair (m',n') with $m' \leq m$ and $n' \leq n$ is also in T.

Suppose that T is an initial segment in $\mathbb{N}^2$, let $\mathbb{K}$ be a field and x, y be variables over $\mathbb{K}$. The elements in $\mathbb{K}[x,y]$ supported on $\mathbb{N}^2 - \mathsf{T}$ form a monomial ideal $I \subset \mathbb{K}[x,y]$. Conversely, any initial segment T in $\mathbb{N}^2$ can be obtained in this manner from a monomial ideal I, as the set of exponents of monomials not in I. If T is finite, we write the minimal monomial generators of I as

$$\boldsymbol{E} = (y^{n_0}, x^{m_1} y^{n_1}, \ldots, x^{m_{s-1}} y^{n_{s-1}}, x^{m_s})$$

with the m_i's increasing and the n_i's decreasing, and we set $m_0 = n_s = 0$. We call n_0 the *height* of T and m_s its *width.* We say that T is *determined* by I, or equivalently by $\boldsymbol{E}$.

For $i = 1, \dots, s$, we set $d_i = m_i - m_{i-1}$, so that $m_i = d_1 + \cdots + d_i$. Then, the cardinal δ of T can be written as $\sum_{i=1}^{s} d_i n_{i-1}$; δ is also called the *degree* of $\boldsymbol{E}$. Similarly, for $i = 1, \dots, s$, we write $e_i = n_{i-1} - n_i$. These definitions are illustrated in Figure 4.1, where the monomials in $\boldsymbol{E}$ are the initial terms of the Gröbner basis in Eq. (4.1.0.1).

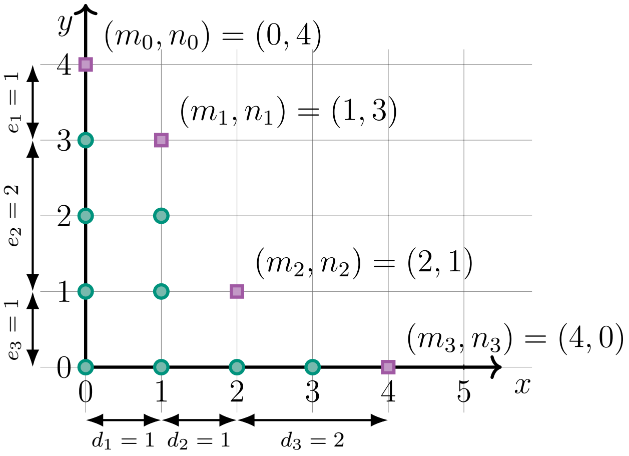


Figure 4.1: An initial segment T (green) and the monomials $\boldsymbol{E} = (y^4, xy^3, x^2y, x^4)$ (purple), with $s = 3$ and $\delta = 9$.

The cost analyses in this paper will be done using in particular the parameters s and δ. If desired, one can simplify such expressions using the following explicit upper bound for s.

Lemma 4.2.1. *The integer s is in $O(\sqrt{\delta})$, and this bound is sharp in some instances.*

Proof. Start from the equality $\delta = \sum_{i=1}^{s} d_i n_{i-1}$, which implies $\delta \geq \sum_{i=1}^{s} n_{i-1}$. Since $n_s = 0$ and $n_{i-1} > n_i$, we get by induction $n_i \geq s - i$ for all i. This implies $\delta \geq s(s-1)/2$, so that s is in $O(\sqrt{\delta})$. For the lower bound, for any integer d we can take $\boldsymbol{E} = (x^i y^{d-i},\ i = 0, \dots, d)$, for which $s = d$ and $\delta = d(d+1)/2$. □

4.2.1.2 Translates of an initial segment We will occasionally make use of the following construction. Let T be a finite initial segment in $\mathbb{N}^2$, and suppose that T is determined by a monomial ideal I, with minimal monomial generators $\boldsymbol{E}$ as above.

For $i = 0, \ldots, s$ we let $\mathsf{T}_{\leftarrow i}$ be the initial segment determined by the colon ideal $I : x^{m_i}$, with minimal monomial generators

$$\boldsymbol{E}_{\leftarrow i} = (y^{n_i}, x^{m_{i+1}-m_i} y^{n_{i+1}}, \ldots, x^{m_{s-1}-m_i} y^{n_{s-1}}, x^{m_s - m_i}).$$

The set $\mathsf{T}_{\leftarrow i}$ has height n_i and width $m_s - m_i$; its cardinal will be written δ_i, and is equal to $\sum_{j=i+1}^{s} d_j n_{j-1}$. We call $\mathsf{T}_{\leftarrow i}$ the ith *translate* of T.

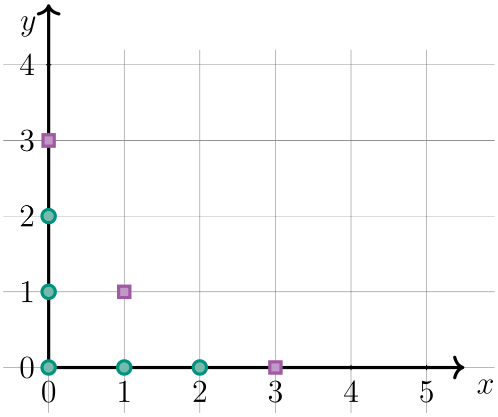


Figure 4.2: The first translate $\mathsf{T}_{\leftarrow 1}$ of T from Figure 4.1.

4.2.1.3 The shell of an initial segment Let T be a finite initial segment in $\mathbb{N}^2$. In this paragraph, we define its *shell* T', which is another initial segment that forms an outer approximation of T with few generators. The definition and the lemma below are from Section 3.4.5 [94, **A.2**].

As we did before, we let

$$\boldsymbol{E} = (y^{n_0}, x^{m_1} y^{n_1}, \ldots, x^{m_{s-1}} y^{n_{s-1}}, x^{m_s})$$

be the minimal monomial generating set associated to T. We define T' by introducing indices $i_\sigma < i_{\sigma-1} < \cdots < i_0$, defined as follows. Set $i_0 = s$. We let $i_1 \geq 0$ be the largest index less than i_0 and such that $m_{i_1} < m_{i_0}/2$, and iterate the process to define a sequence $i_\sigma = 0 < i_{\sigma-1} < \cdots < i_0 = s$. We can then consider the monomials

$$\boldsymbol{E}' = (y^{n_{i_\sigma}}, x^{m_{i_{\sigma-1}}} y^{n_{i_{\sigma-1}}}, \ldots, x^{m_{i_0}}) = (y^{n_0}, x^{m_{i_{\sigma-1}}} y^{n_{i_{\sigma-1}}}, \ldots, x^{m_s}),$$

and let T' be the initial segment determined by $\boldsymbol{E}'$.

Lemma 4.2.2. *The initial segment T' contains T, its cardinal is at most 2δ and σ is in $O(\log(\delta))$.*

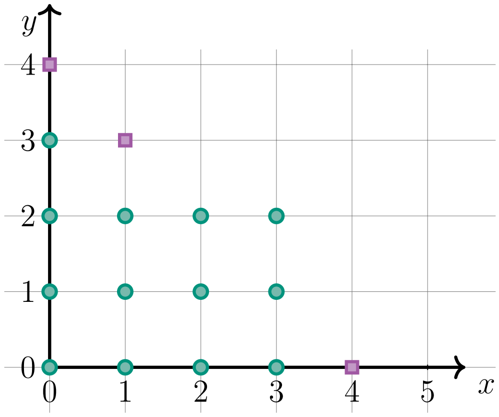


Figure 4.3: The shell of T from Figure 4.1.

In our pseudo-code, we will write $\mathsf{T}' \leftarrow \textsc{Shell}(\mathsf{T})$ to indicate that T' is the shell of T. The algorithm Shell does not use any base field or base ring operation, only index manipulations (in particular, it does not show up in our cost analyses).

4.2.2 Structured polynomial multiplication

We now prove two propositions regarding polynomial multiplication in $\mathbb{A}[x, y]$, for an arbitrary ring $\mathbb{A}$, which will be the basis of the runtime analysis of several algorithms. We mention in all propositions below that the algorithms in this section only use additions and multiplications in $\mathbb{A}$, as we will need this property in the sequel. In what follows, given two sets S, T in $\mathbb{N}^2$, $\mathsf{S} + \mathsf{T}$ denotes their Minkowski sum.

The main prerequisite is the following fact: if $\mathsf{S} \subset \mathbb{N}^2$ is a rectangle, given A and B in $\mathbb{A}[x, y]_{\mathsf{S}}$, we can compute $AB \in \mathbb{A}[x, y]_{\mathsf{S}+\mathsf{S}}$ using $O\tilde{~}(|\mathsf{S}|)$ operations $(+, \times)$ in $\mathbb{A}$: if S contains the origin, this is done using Kronecker substitution to reduce to multiplication in $\mathbb{A}[x]$, see [73, Corollary 8.28]; in the general case, we reduce to the situation where S contains the origin by factoring out $x^u y^v$ from A and B, with (u, v) being the unique minimal element of S.

This being said, the first result we highlight here gives the cost of computing the product AB, for A and B supported on the same initial segment T. Note that AB is supported on $\mathsf{T} + \mathsf{T}$, and that if T has height n and width m, $\mathsf{T} + \mathsf{T}$ has cardinal $\Theta(nm)$. Indeed, this set contains the rectangle $\{0, \dots, m-1\} \times \{0, \dots, n-1\}$ of cardinal nm, and is contained in the rectangle $\{0, \dots, 2m-2\} \times \{0, \dots, 2n-2\}$ of cardinal less than $4nm$, so that $|\mathsf{T} + \mathsf{T}| \in \Theta(nm)$. This is to be contrasted with the cardinal of T itself, which can range anywhere between $n + m$ and nm.

Proposition 4.2.1. *Consider a finite initial segment* $\mathsf{T} \subset \mathbb{N}^2$*, of height* n *and width* m*. Given* A *and* B *in* $\mathbb{A}[x,y]_{\mathsf{T}}$*, one can compute* AB *using* $O\tilde{}(|\mathsf{T}+\mathsf{T}|) = O\tilde{}(nm)$ *operations* $(+,\times)$ *in* $\mathbb{A}$*.*

Proof. Let S be the rectangle $\{0,\dots,m-1\} \times \{0,\dots,n-1\}$, so that S contains T. Then, A and B are in $\mathbb{A}[x,y]_{\mathsf{S}}$, so we can multiply them using $O\tilde{}(|\mathsf{S}+\mathsf{S}|) = O\tilde{}(nm)$ operations $(+,\times)$ in $\mathbb{A}$ with Kronecker substitution, as pointed out above, and this runtime is also $O\tilde{}(|\mathsf{T}+\mathsf{T}|)$. □

Our second proposition gives an algorithm to compute $AB \in \mathbb{A}[x,y]$, where A is supported on a rectangle containing the origin and B on an initial segment.

Proposition 4.2.2. *Consider a rectangle* $\mathsf{S} \subset \mathbb{N}^2$ *and a finite initial segment* $\mathsf{T} \subset \mathbb{N}^2$*. Given* A *in* $\mathbb{A}[x,y]_{\mathsf{S}}$ *and* B *in* $\mathbb{A}[x,y]_{\mathsf{T}}$*, one can compute* AB *using* $O\tilde{}(|\mathsf{S}+\mathsf{T}|)$ *operations* $(+,\times)$ *in* $\mathbb{A}$*.*

Without loss of generality, we assume that S contains the origin $(0,0)$; if not, as above, factor out the monomial $x^u y^v$ from A, with (u,v) the minimal element in S. We can thus suppose that S is the rectangle $\{0,\dots,\ell-1\} \times \{0,\dots,h-1\}$, for some integers $\ell, h \geq 1$, so in particular $|\mathsf{S}| = \ell h$, and that T is an initial segment of cardinal $|\mathsf{T}| = \delta$, with height n and width m.

If $\mathbb{A}$ is a field of characteristic zero, this result follows directly from the sparse evaluation and interpolation algorithms of [36]. More generally, if $\mathbb{A}$ is a field of cardinal at least $\max(\ell+m, h+n)-1$, this is also the case, using the algorithm in [97]. The algorithm below achieves the same asymptotic runtime, without assumption on $\mathbb{A}$. The proof is slightly more involved than that of the previous proposition, and occupies the rest of this section.

4.2.2.1 An algorithm when T is a rectangle Suppose first that $\mathsf{T} = \{0,\dots,m-1\} \times \{0,\dots,n-1\}$, so that $\delta = nm$; then the cardinal of $\mathsf{S}+\mathsf{T}$ is $(\ell+m-1)(h+n-1)$.

Take A in $\mathbb{A}[x,y]_{\mathsf{S}}$ and B in $\mathbb{A}[x,y]_{\mathsf{T}}$. Then, both A and B are in $\mathbb{A}[x,y]_{\mathsf{S}+\mathsf{T}}$. Since $\mathsf{S}+\mathsf{T}$ is a rectangle, we saw in the preamble of this section that using Kronecker's substitution, we can compute their product using $O\tilde{}(|\mathsf{S}+\mathsf{T}|) = O\tilde{}((\ell+m-1)(h+n-1))$ operations $(+,\times)$ in $\mathbb{A}$. In the main algorithm below, this is written KroneckerMultiply(A,B).

4.2.2.2 A first general algorithm We now suppose that T is an arbitrary initial segment, and that it is determined by the monomials

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s}),$$

with the m_i's increasing, the n_i's decreasing, and $m_0 = n_s = 0$; note that we also have $n_0 = n$ and $m_s = m$. As before, for $i = 1, \dots, s$, we set $d_i = m_i - m_{i-1}$, so that $m_i = d_1 + \cdots + d_i$.

The input $B \in \mathbb{A}[x, y]_{\mathsf{T}}$ can then be written as $B = \sum_{0 \le i < s} B_i x^{m_i}$, with B_i supported on $\mathsf{T}_i = \{0, \dots, d_{i+1} - 1\} \times \{0, \dots, n_i - 1\}$. To compute AB, with A in $\mathbb{A}[x, y]_{\mathsf{S}}$, we thus compute all AB_i and add up the results.

Algorithm 4.2.1 MultiplyNaive($A, \mathsf{S}, B, \mathsf{T}$)

```
Input: $A$ in $\mathbb{A}[x,y]_{\mathsf{S}}$, $B$ in $\mathbb{A}[x,y]_{\mathsf{T}}$
Output: $AB$ in $\mathbb{A}[x,y]_{\mathsf{S}+\mathsf{T}}$
write $B = B_0 + B_1 x^{m_1} + \cdots + B_{s-1} x^{m_{s-1}}$ with $B_i \in \mathbb{A}[x,y]_{\{0,\dots,d_{i+1}-1\}\times\{0,\dots,n_i-1\}}$
    for all $i$
for $i = 0, \dots, s-1$ do $C_i \leftarrow$ KroneckerMultiply($A, B_i$)
return $C_0 + C_1 x^{m_1} + \cdots + C_{s-1} x^{m_{s-1}}$
```

By the result in the previous paragraph, each product AB_i can be computed in

$$O\tilde{}((\ell + d_{i+1} - 1)(h + n_i - 1)) = O\tilde{}((\ell - 1)(h - 1) + (\ell - 1)n_i + d_{i+1}(h - 1) + d_{i+1}n_i)$$

operations in $\mathbb{A}$, and the cost of adding this product to the final result fits into the same bound. Using the inequality $n_i \le n_0 = n$ for all i, as well as $d_1 + \cdots + d_s = m_s = m$ and $d_1 n_0 + \cdots + d_s n_{s-1} = \delta$ (the cardinal of T), we see that the total cost is

$$O\tilde{}(s(\ell - 1)(h - 1) + s(\ell - 1)n + m(h - 1) + \delta).$$

On the other hand, we can determine the cardinal of the sum $\mathsf{U} = \mathsf{S} + \mathsf{T}$ as follows. The set U is the disjoint union of the following sets:

- $\mathsf{U}_1 = \{0, \dots, \ell - 2\} \times \{0, \dots, h - 2\}$,
- $\mathsf{U}_2 = (0, h - 1) + \{0, \dots, \ell - 2\} \times \{0, \dots, n - 1\}$
- $\mathsf{U}_3 = (\ell - 1, 0) + \{0, \dots, m - 1\} \times \{0, \dots, h - 2\}$
- $\mathsf{U}_4 = (\ell - 1, h - 1) + \mathsf{T}$.

This is established by taking (i, j) in S, (v, w) in T, and discussing according to the signs of $v - (\ell - 1 - i)$ and $w - (h - 1 - j)$. As a result, we obtain

$$|\mathsf{S} + \mathsf{T}| = (\ell - 1)(h - 1) + (\ell - 1)n + m(h - 1) + \delta.$$

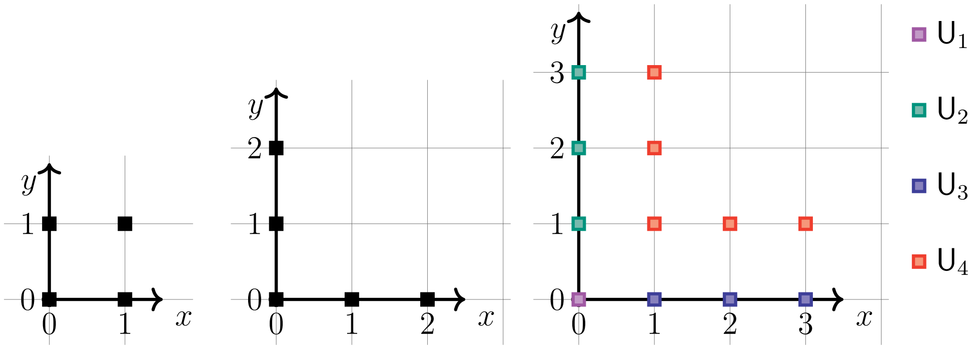


Figure 4.4: The sets S, T and $\mathsf{U} = \mathsf{S} + \mathsf{T}$, with $\ell = h = 2$ and $n = m = 3$.

4.2.2.3 The main algorithm. The runtime reported above does not fit in the target cost $O\tilde{}(|\mathsf{S} + \mathsf{T}|)$, as s could be large. To circumvent this issue, we apply the algorithm of the previous paragraph, but we replace T by its shell T'. We know (Lemma 4.2.2) that the cardinal of T' is at most 2δ, that its width and height are the same as those of T, and that it is generated by $\sigma \in O(\log(s)) \subset O(\log(\delta))$ terms.

Algorithm 4.2.2 Multiply$(A, \mathsf{S}, B, \mathsf{T})$

```
Input: $A$ in $\mathbb{A}[x,y]_\mathsf{S}$, $B$ in $\mathbb{A}[x,y]_\mathsf{T}$
Output: $AB$ in $\mathbb{A}[x,y]_{\mathsf{S}+\mathsf{T}}$
$\mathsf{T}' \leftarrow$ Shell$(\mathsf{T})$
return MultiplyNaive$(A, \mathsf{S}, B, \mathsf{T}')$
```

The algorithm of the previous paragraph still applies (since T is contained in T'), and its runtime is then $O\tilde{}((\ell-1)(h-1)\log(\delta)+(\ell-1)n\log(\delta)+m(h-1)+\delta)$ operations $(+, \times)$ in $\mathbb{A}$. Since we saw that $|\mathsf{S}+\mathsf{T}| = (\ell-1)(h-1)+(\ell-1)n+m(h-1)+\delta$, the above expression is indeed in $O\tilde{}(|\mathsf{S}+\mathsf{T}|)$. This finishes the proof of Proposition 4.2.2.

4.3 Lexicographic Gröbner bases

In this section, we first review Lazard's structure theorem [109] for lexicographic Gröbner bases in $\mathbb{K}[x, y]$, for a field $\mathbb{K}$, then a parametrization of such bases due to [40]. While the core of the discussion makes no assumption on the ideals we consider, we also highlight the case of ideals that are primary at the origin, that is, $\langle x, y\rangle$-primary.

In all that follows, we use the lexicographic monomial order $\succ$ on $\mathbb{K}[x,y]$ induced by $y \succ x$.

4.3.1 The structure theorem

Consider a zero dimensional ideal $I \subseteq \mathbb{K}[x,y]$, and let $\mathcal{G} = (g_0, \ldots, g_s)$ be its reduced minimal Gröbner basis, listed in decreasing order. Let further

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$$

be the minimal reduced basis of the initial ideal $in(I)$ of I, listed in decreasing order, so the n_i's are decreasing and the m_i's are increasing; as before, we set $m_0 = n_s = 0$.

It follows that g_i has initial term $x^{m_i}y^{n_i}$ for all i; in particular g_0 is monic in y with initial term y^{n_0}.

As in Section 4.2.1, for $i = 1, \ldots, s$, we set $d_i = m_i - m_{i-1}$, with thus $m_i = d_1 + \cdots + d_i$, and $e_i = n_{i-1} - n_i$.

Lazard proved in [109, Theorem 1] the existence of polynomials $D_1, \ldots, D_s$ in $\mathbb{K}[x]$, all monic in x and of respective degrees $d_1, \ldots, d_s$, such that for $i = 0, \ldots, s$, g_i can be written as M_iG_i, with $M_i = D_1 \cdots D_i \in \mathbb{K}[x]$ and $G_i \in \mathbb{K}[x,y]$ monic of degree n_i in y (for $i = 0$, we set $D_0 = 1$). In particular, for $i = s$, this gives $g_s = M_s = D_1 \cdots D_s$ and $G_s = 1$. In addition, for $i = 0, \ldots, s-1$, we have the membership relation

$$G_i \in \langle G_{i+1},\ D_{i+2}G_{i+2},\ \ldots,\ D_{i+2} \cdots D_s \rangle = \left\langle \frac{g_{i+1}}{M_{i+1}}, \frac{g_{i+2}}{M_{i+1}}, \ldots, \frac{g_s}{M_{i+1}} \right\rangle, \qquad (4.3.0.1)$$

where the polynomials $G_{i+1}, D_{i+2}G_{i+2}, \ldots, D_{i+2} \cdots D_s$ also form a zero-dimensional Gröbner basis.

If $\mathcal{G}$ generates an $\langle x,y \rangle$-primary ideal, we have $D_i = x^{m_i}$ for all i, with thus $g_s = x^{m_s}$. Besides, for all i, $G_i(0,y)$ vanishes only at $y = 0$, *i.e.* $G_i(0,y) = y^{n_i}$, see [109, Theorem 2].

In terms of data structures, representing $\mathcal{G} = (g_0, \ldots, g_s)$ involves $O(s\delta)$ field elements, with δ the degree of I. As a remark, we note that it would be sufficient to store the polynomials $\boldsymbol{D} = (D_1, \ldots, D_s)$ and $\boldsymbol{G} = (G_0, \ldots, G_s)$ instead. If $\mathsf{T} \subset \mathbb{N}^2$ is the initial segment determined by $\boldsymbol{E}$, the structure theorem implies that for $i = 0, \ldots, s$, $G_i - y^{n_i}$ is supported on the ith translate $\mathsf{T}_{\leftarrow i}$ of T. In particular, δ_i field elements are needed to store it, with $\delta_i = |\mathsf{T}_{\leftarrow i}|$, hence a slightly improved total of $O(\sum_{i=0}^{s} \delta_i)$ field elements for $\boldsymbol{D}$ and $\boldsymbol{G}$.

4.3.2 Conca and Valla's parametrization

In this subsection, we suppose that the tuple $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$ is fixed. Following [40], we are interested in describing the set of ideals I in $\mathbb{K}[x, y]$ that have initial ideal generated by $\boldsymbol{E}$. We call this set the *Gröbner cell* of $\boldsymbol{E}$, and we write it $\mathcal{C}(\boldsymbol{E}) := \{I \mid in(I) = \langle \boldsymbol{E} \rangle\}$. We will also mention a subset of it, the set of ideals I in $\mathbb{K}[x, y]$ with initial ideal generated by $\boldsymbol{E}$ and that are $\langle x, y\rangle$-primary; this is called the *punctual* Gröbner cell of $\boldsymbol{E}$, and is written $\mathcal{C}_0(\boldsymbol{E})$.

The idea of describing ideals with a prescribed initial ideal goes back to [33, 32, 95] for ideals in $\mathbb{K}[\![x, y]\!]$ and [69] for $\mathbb{K}[x_1, \dots, x_n]$; it was then developed in [131, 139, 113] and several further references. It is known that these Gröbner cells, also called strata, have corresponding moduli spaces that are affine spaces, but to our knowledge, no general explicit description has yet been given. In our case however, Conca and Valla obtained in [40] a complete description of Gröbner cells and punctual Gröbner cells for bivariate ideals under the lexicographic order (following previous work of [65], where the dimensions of these cells were already made explicit).

❧ **Example 4.3.1**

For an example of a punctual Gröbner cell, taking $\boldsymbol{E} = (y^4, xy^3, x^2y, x^4)$ as in Figure 4.1, using the facts that $g_i = x^{m_i}G_i$ and that $G_i(0, y) = y^{n_i}$, we deduce that the lexicographic Gröbner basis of an ideal in $\mathcal{C}_0(\boldsymbol{E})$ necessarily has the following shape, for some coefficients $c_1, \dots, c_8$ in $\mathbb{K}$:

$$\begin{aligned}
g_1 &= y^4 + c_1xy^2 + c_2xy + c_3x^3 + c_4x^2 + c_5x \\
g_2 &= xy^3 + c_6x^3 + c_7x^2 \\
g_3 &= x^2y + c_8x^3 \\
g_4 &= x^4
\end{aligned}$$

So far, though, we have not taken into account the membership equality in (4.3.0.1), which imposes relations on the coefficients c_i. The parametrizations of $\mathcal{C}(\boldsymbol{E})$ and $\mathcal{C}_0(\boldsymbol{E})$ given below resolve this issue.

Recall that we write $d_i = m_i - m_{i-1}$ and $e_i = n_{i-1} - n_i$, for $i = 1, \dots, s$. Given I in $\mathcal{C}(\boldsymbol{E})$, Conca and Valla prove the existence and uniqueness of polynomials $(\sigma_{j,i})_{0 \le i \le s-1, i \le j \le s}$ in $\mathbb{K}[x, y]$ with the following degree constraints:

- for all $i = 0, \dots, s-1$ and $j = i, \dots, s$, $\deg(\sigma_{j,i}, x) < d_{i+1}$

- for all $i = 0, \dots, s-1$, $\sigma_{i,i}$ is in $\mathbb{K}[x]$ and $\deg(\sigma_{j,i}, y) < e_j$ holds for $j = i+1, \dots, s$,

and such that the following properties hold. Define polynomials $\mathcal{H} = (h_0, \dots, h_s)$ in $\mathbb{K}[x, y]$ by

- $h_s = (x^{d_1} - \sigma_{0,0}) \cdots (x^{d_s} - \sigma_{s-1,s-1})$
- for $i = 0, \dots, s-1$,

$$x^{d_{i+1}} h_i - y^{e_{i+1}} h_{i+1} = \sigma_{i,i} h_i + \sigma_{i+1,i} h_{i+1} + \cdots + \sigma_{s,i} h_s; \tag{4.3.0.2}$$

then, all polynomials h_i's are in I. Since the relations above imply that for $i = 0, \dots, s$, h_i has initial term $x^{m_i} y^{n_i}$, $\mathcal{H} = (h_0, \dots, h_s)$ is a minimal Gröbner basis of I. (Note that Eq. (4.3.0.2) then gives the normal form of the syzygy between h_i and h_{i+1}.)

Conversely, for any choice of the polynomials $\sigma_{j,i}$ satisfying the degree constraints above, the resulting polynomials $\mathcal{H}$ form a minimal Gröbner basis of an ideal I in $\mathcal{C}(\boldsymbol{E})$.

Let us briefly mention some properties of the polynomials $h_0, \dots, h_s$. First, we claim that they have x-degree either exactly m_s (for h_s), or less than m_s, for $h_0, \dots, h_{s-1}$. This is true for h_s by construction. For the other indices, this follows from a decreasing induction, by rewriting (4.3.0.2) as

$$(x^{d_{i+1}} - \sigma_{i,i}) h_i = y^{e_{i+1}} h_{i+1} + \sigma_{i+1,i} h_{i+1} + \cdots + \sigma_{s,i} h_s, \tag{4.3.0.3}$$

where all terms $\sigma_{j,i} h_j$ on the right have x-degree less than $d_{i+1} + m_s$.

Next, note that for $i = 0, \dots, s$, $(x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1})$ divides h_i, and thus all polynomials $h_i, \dots, h_s$; this follows from (4.3.0.3) by a decreasing induction (for $i = 0$, the empty product is set to 1). Since h_i has initial term $x^{m_i} y^{n_i} = x^{d_1 + \cdots + d_i} y^{n_i}$, we deduce that $(x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1})$ is precisely the polynomial coefficient of y^{n_i} in h_i.

Let then $\mathcal{G} = (g_0, \dots, g_s)$ be the reduced Gröbner basis obtained by inter-reducing $\mathcal{H}$. Since none of the terms in $(x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1}) y^{n_i}$ can be reduced by $h_0, \dots, h_{i-1}$ or $h_{i+1}, \dots, h_s$, we see that $(x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1})$ is also the polynomial coefficient of y^{n_i} in g_i. Hence, the polynomials D_i and M_i that appear in Lazard's structure theorem are respectively given by $D_i = x^{d_i} - \sigma_{i-1,i-1}$ and $M_i = (x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1})$.

Altogether, the total number N of coefficients that appear in the polynomials $(\sigma_{j,i})_{0\le i\le s-1, i\le j\le s}$, for the Gröbner cell $\mathcal{C}(\boldsymbol{E})$, is given by

$$\begin{aligned} N &= \sum_{i=0}^{s-1}\left(\sum_{j=i+1}^{s} d_{i+1}e_j + d_{i+1}\right) \\ &= \sum_{i=0}^{s-1} d_{i+1}n_i + \sum_{i=0}^{s-1} d_{i+1} \\ &= \delta + m_s, \end{aligned}$$

with δ the degree of $\boldsymbol{E}$. These coefficients will be written $\lambda_1, \ldots, \lambda_N$ and called *Gröbner parameters*; this gives us a bijection $\Phi_{\boldsymbol{E}}$ between $\mathbb{K}^N$ and $\mathcal{C}(\boldsymbol{E})$.

The elements in the *punctual* Gröbner cell $\mathcal{C}_0(\boldsymbol{E})$ are obtained by setting some of the Gröbner parameters to zero, corresponding to the following extra conditions:

- the polynomials $\sigma_{0,0}, \ldots, \sigma_{s-1,s-1}$ vanish (recall that for the punctual Gröbner cell, we have $D_i = x^{d_i}$ and $M_i = x^{m_i}$ for all i)
- $\sigma_{i+1,i}$ is divisible by x, for $i = 0, \ldots, s-1$.

The number of remaining coefficients in $\sigma_{1,0}, \ldots, \sigma_{s,s-1}$ is

$$\begin{aligned} N_0 &= \sum_{i=0}^{s-1}\left(\sum_{j=i+1}^{s} d_{i+1}e_j - e_{i+1}\right) \\ &= \sum_{i=0}^{s-1} d_{i+1}n_i - \sum_{i=0}^{s-1} e_{i+1} \\ &= \delta - n_0, \end{aligned}$$

establishing a bijection between $\mathbb{K}^{N_0}$ and $\mathcal{C}_0(\boldsymbol{E})$. Recall that in the primary case, the degree δ of $\boldsymbol{E}$ is the multiplicity of the ideals in $\mathcal{C}_0(\boldsymbol{E})$ at the origin.

❧ Example 4.3.2

Let us describe the punctual Gröbner cell of $\boldsymbol{E}$ in our running example (Example 4.1.1). It has dimension $N_0 = 9 - 4 = 5$, so that we can use parameters $\lambda_1, \ldots, \lambda_5$, with polynomials $(\sigma_{i,j})$ of the form

$$\sigma_{0,0} = \sigma_{1,0} = 0, \quad \sigma_{2,0} = \lambda_1 y + \lambda_2, \quad \sigma_{3,0} = \lambda_3,$$

$$\sigma_{1,1} = n_{2,1} = 0, \quad \sigma_{3,1} = \lambda_4, \quad \sigma_{2,2} = 0, \quad \sigma_{3,2} = \lambda_5 x.$$

Then, the ideals in $\mathcal{C}_0(\boldsymbol{E})$ are exactly those ideals with Gröbner bases as follows:

$$\begin{aligned}
h_0 &= y^4 + \lambda_5 xy^3 + \lambda_1 xy^2 + (\lambda_1\lambda_5 + \lambda_4)x^2y + \lambda_2 xy + \lambda_3 x^3 + \lambda_2\lambda_5 x^2\\
h_1 &= xy^3 + \lambda_5 x^2y^2 + \lambda_4 x^3\\
h_2 &= x^2y + \lambda_5 x^3\\
h_3 &= x^4.
\end{aligned}$$

As expected, these are not reduced Gröbner bases. After reduction, we obtain the following polynomials $\mathcal{G}$:

$$\begin{aligned}
g_0 &= y^4 + \lambda_1 xy^2 + \lambda_2 xy + (-\lambda_1\lambda_5^2 + \lambda_3 - 2\lambda_4\lambda_5)x^3 + \lambda_2\lambda_5 x^2\\
g_1 &= xy^3 + \lambda_4 x^3\\
g_2 &= x^2y + \lambda_5 x^3\\
g_3 &= x^4.
\end{aligned} \tag{4.3.0.4}$$

4.4 Reduction modulo a lexicographic Gröbner basis

As before, suppose that $\mathcal{G} = (g_0, \dots, g_s)$ is a lexicographic Gröbner basis in $\mathbb{K}[x, y]$, with initial segment $\mathsf{T} \subset \mathbb{N}^2$. Given f in $\mathbb{K}[x, y]$, we are interested in computing the remainder $r = f \text{ rem } \mathcal{G} \in \mathbb{K}[x, y]_{\mathsf{T}}$; this will be used on multiple occasions in this paper, and is also an interesting question in itself. Remarkably, we are not aware of previous work on the complexity of this particular question.

We start by developing the necessary background as a problem in plane geometry. This is inspired by work of [90], which was specific to certain weighted orderings (we discuss this further below). We continue with algorithms to convert polynomials into a so-called *mixed-radix* representation, and back; the reduction algorithm itself is then given in the last subsection.

4.4.1 A paving problem

For $\mathcal{G}$ as above and f in $\mathbb{K}[x, y]$, the remainder $r = f \text{ rem } \mathcal{G}$ is uniquely defined, but the quotients Q_i in the relation $f = Q_0 g_0 + \cdots + Q_s g_s + r$ are not. The reduction algorithm will obtain r by computing the Q_i's one after the other. Hence, to completely specify the algorithm, we need to make these quotients unambiguous: whenever a monomial $x^u y^v$ can be reduced by more than one of the Gröbner basis elements, we must prescribe which of the g_i's is used. The cost of the resulting algorithm will depend in an essential manner on these decisions.

[90] introduced a dichotomic scheme, in the context of reduction modulo certain "nice" Gröbner bases, for weighted degree orderings. In this subsection, we adapt their construction to our situation.

As before, suppose that the initial terms of $\mathcal{G}$ are the monomials

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s});$$

we still write $d_i = m_i - m_{i-1}$ and $e_i = n_{i-1} - n_i$, for $i = 1, \ldots, s$. The set of monomials to which we will apply the main reduction algorithm is $\{x^u y^v,\ 0 \leq u < m_s,\ 0 \leq v < n_0\}$, so it has cardinal $n_0 m_s$ (the general case will be reduced to this situation). In particular, neither g_0 nor g_s can reduce any of these monomials.

We can then translate our question into a paving problem in the plane. We want to cover $\mathsf{S} = \{0, \ldots, m_s - 1\} \times \{0, \ldots, n_0 - 1\} - \mathsf{T}$ by rectangles, under the following constraints:

- we use $s-1$ pairwise disjoint rectangles, $\mathsf{R}_1, \ldots, \mathsf{R}_{s-1}$, so that R_i will index the set of monomials that are reduced using g_i
- for all i, R_i has the form $\{m_i, \ldots, m_{i+\ell_i} - 1\} \times \{n_i, \ldots, n_{i-h_i} - 1\}$, for some positive integers ℓ_i, h_i such that $i + \ell_i \leq s$ and $i - h_i \geq 0$
- the union of all R_i's covers S.

The sequence $((\ell_1, h_1), \ldots, (\ell_{s-1}, h_{s-1}))$ is sufficient to specify such a paving. Our goal is then to minimize the quantity

$$c := n_0 \sum_{i=1}^{s-1} (m_{i+\ell_i} - m_i) + m_s \sum_{i=1}^{s-1} (n_{i-h_i} - n_i),$$

where $(m_{i+\ell_i} - m_i)$ and $(n_{i-h_i} - n_i)$ are respectively the width and height of R_i. This quantity will turn out to determine the cost of the reduction algorithm; the target is to keep c in $O\tilde{}(n_0 m_s)$, since we mentioned that $n_0 m_s$ in an upper bound on the number of monomials in the polynomials we want to reduce.

The following figure shows two possible pavings, for the case $d = 4$ of the family already seen in the proof of Lemma 4.2.1, with $\boldsymbol{E} = (y^d, xy^{d-1}, \ldots, x^d)$. For this family, $n_0 = m_s = d$ and $n_0 m_s = d^2$; the strategies showed on the example below have either $\sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i)$ or $\sum_{i=1}^{s-1}(n_{i-h_i} - n_i)$ in $\Theta(d^2)$, so c is in $\Theta(d^3) = \Theta((n_0 m_s)^{1.5})$ in either case.

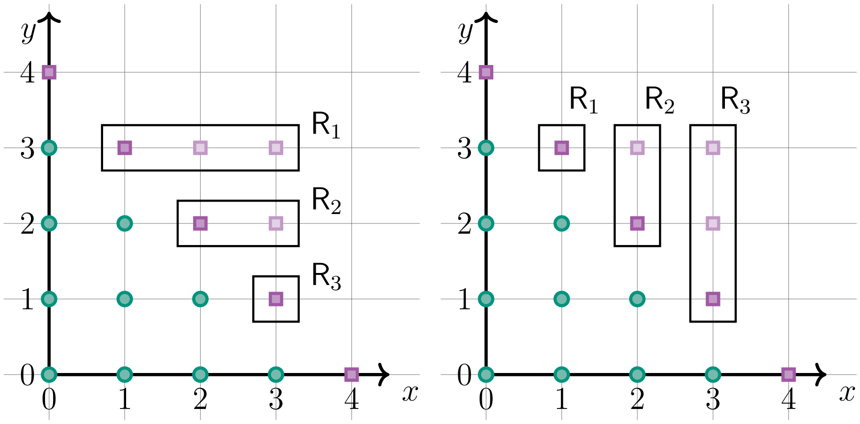


Figure 4.5: Two possible pavings with $d = 4$.

For this family, a better solution is given below.

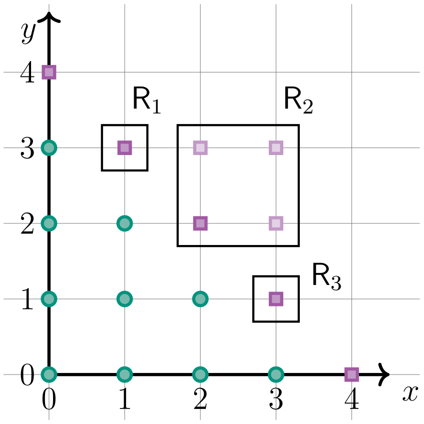


Figure 4.6: An improved paving.

This design was introduced in [90], for families $\boldsymbol{E}$ similar to the one in the example, where the step widths d_i are (almost) constant, and all step heights e_i are equal to 1. The construction we give below for arbitrary inputs is derived from it in a direct manner. In what follows, $\mathrm{val}_2(i)$ denotes the 2-adic valuation of a positive integer i.

❧ **Definition 27.** *For $i = 1, \ldots, s-1$, define:*

- $h_i = 2^{\mathrm{val}_2(i)}$
- $\ell_i = \min(h_i, s - i)$

As a result, the rectangle R_i is $\{m_i, \ldots, m_{\min(i+h_i,s)} - 1\} \times \{n_i, \ldots, n_{i-h_i} - 1\}$.

Proposition 4.4.1. *For any s and any choices of $m_1, \ldots, m_s$ and $n_0, \ldots, n_{s-1}$, the rectangles $\mathsf{R}_1, \ldots, \mathsf{R}_{s-1}$ are pairwise disjoint, cover $\mathsf{S} = \{0, \ldots, m_s - 1\} \times \{0, \ldots, n_0 - 1\} - \mathsf{T}$, and satisfy $i + \ell_i \leq s$ and $i - h_i \geq 0$ for all i.*

Proof. The last claim is a direct consequence of the definitions. We prove the rest of the proposition by reduction to the case where all d_i's and e_i's are equal to one. The proof is technical but raises no special difficulty.

For any positive integer s, we define the monomials $\mathscr{E}_s = (x^i y^{s-i},\ 0 \leq i \leq s)$, the initial segment $\mathscr{T}_s$ determined by $\mathscr{E}_s$ and $\mathscr{S}_s = \{0, \ldots, s-1\} \times \{0, \ldots, s-1\} - \mathscr{T}_s$; note that $\mathscr{T}_s$ is the set of all pairs of non-negative integers (a, b) with $b < s - a$. Finally, for $i = 1, \ldots, s-1$ we define the rectangle $\mathscr{R}_{i,s} = \{i, \ldots, \min(i + h_i, s) - 1\} \times \{s - i, \ldots, s - i + h_i - 1\} \subset \mathscr{S}_s$.

We start from $m_1, \ldots, m_s$ and $n_0, \ldots, n_{s-1}$ as in the proposition's statement, with corresponding sets T and S in $\mathbb{N}^2$. Take a point (u, v) in S. Because $u < m_s$, there exists a unique pair (α, u') such that $u = m_\alpha + u'$, with $0 \leq \alpha \leq s - 1$ and $0 \leq u' < d_{\alpha+1}$. Similarly, because $v < n_0$, there exists a unique pair (β, v') such that $v = n_\beta + v'$, with $1 \leq \beta \leq s$ and $0 \leq v' < e_\beta$. We claim that $(\alpha, s - \beta)$ is in the set $\mathscr{S}_s$ defined in the previous paragraph, and that for $i = 1, \ldots, s - 1$, (u, v) is in the rectangle R_i if and only if $(\alpha, s - \beta)$ is in the rectangle $\mathscr{R}_{i,s}$.

- For the first claim, we already pointed out the inequalities $0 \leq \alpha \leq s - 1$ and $1 \leq \beta \leq s$, which gives $0 \leq s - \beta \leq s - 1$, so that $(\alpha, s - \beta)$ is in the square $\{0, \ldots, s-1\} \times \{0, \ldots, s-1\}$. On the other hand, we have $v \geq n_\alpha$ (otherwise (α, β) would be in T), and so $\beta \leq \alpha$ and $s - \beta \geq s - \alpha$. This proves that the point $(\alpha, s - \beta)$ is not in $\mathscr{T}_s$, so altogether, it lies in $\mathscr{S}_s$.

- For the second claim, note that since $u = m_\alpha + u'$, with $0 \leq u' < d_{\alpha+1}$, $m_i \leq u < m_{\min(i+h_i,s)}$ is equivalent to $i \leq \alpha < \min(i + h_i, s)$. Similarly, the inequalities $n_i \leq v < n_{i-h_i}$ are equivalent to $s - i \leq s - \beta < s - i + h_i$. This proves the claim.

To conclude, it is now sufficient to prove that for all s, the following property, written $P(s)$, holds: the rectangles $\mathscr{R}_{1,s}, \ldots, \mathscr{R}_{s-1,s}$ are pairwise disjoint and cover $\mathscr{S}_s$. First, we prove it for s a power of two, of the form $s = 2^k$, by induction on $k \geq 1$. For $k = 1$ (so $s = 2$), there is nothing to prove, as $\mathscr{S}_2 = \{1\} \times \{1\} = \mathscr{R}_{1,2}$.

Supposing that $P(s)$ is true for $s = 2^k$, we now prove it for $s' = 2s$. For $\mathscr{S}$ a subset of $\mathbb{N}^2$, we write $\mathscr{S} \cap \{x \leq t\}$ for the set of all (x, y) in $\mathscr{S}$ with $x \leq t$. The sets $\mathscr{S} \cap \{x \geq t\}$, $\mathscr{S} \cap \{x \leq t, y \leq t'\}$, etc, are defined similarly.

First, we note that for any power of two $\sigma = 2^t$ and $i = 1, \ldots, \sigma - 1$, we have $i + h_i \leq \sigma$, so the rectangle $\mathscr{R}_{i,\sigma}$ is simply $\mathscr{R}_{i,\sigma} = \{i, \ldots, i+h_i-1\} \times \{\sigma - i, \ldots, \sigma - i + h_i - 1\}$. As a result, the rectangles $\mathscr{R}_{1,s'}, \ldots, \mathscr{R}_{s-1,s'}$ are translates of $\mathscr{R}_{1,s}, \ldots, \mathscr{R}_{s-1,s}$ by $(0, s)$, so by the induction assumption, they are pairwise disjoint, cover $\mathscr{S}_{s'} \cap \{x \leq s-1\}$, and do not meet $\mathscr{S}_{s'} \cap \{x \geq s\}$ (on Figure 4.6, we have $s = 2$, $s' = 4$, and there is only one such rectangle, written R_1). Since $h_i = \varepsilon_{i+s}$ for $i = 1, \ldots, s-1$, we also deduce that the rectangles $\mathscr{R}_{s+1,s'}, \ldots, \mathscr{R}_{2s-1,s'}$ are translates of $\mathscr{R}_{1,s}, \ldots, \mathscr{R}_{s-1,s}$ by $(s, 0)$. Thus, they are pairwise disjoint, cover $\mathscr{S}_{s'} \cap \{x \geq s, y \leq s-1\}$, and do not meet $\mathscr{S}_{s'} \cap \{x \geq s, y \geq s\}$ (on Figure 4.6, this is R_3). Finally, $\mathscr{R}_{s,s'}$ is the rectangle $\{s, 2s-1\} \times \{s, 2s-1\}$ (on Figure 4.6, this is R_2). Altogether, $P(s')$ holds and the induction is complete.

The last step is to prove that $P(s)$ holds for all s, knowing that it holds for all powers of two. Let s be arbitrary and let s' be the first power of two greater than or equal to s, so that we know that $P(s')$ holds. Let $s'' = s'/2$. Since $s' < 2s$, $s'' \leq s$. For $i < s''$, $\mathscr{R}_{i,s} = \mathscr{R}_{i,s'} - (s'-s, 0)$, whereas for $s'' \leq i \leq s-1$, $\mathscr{R}_{i,s} = \mathscr{R}_{i,s'} \cap \{x \leq s-1\} - (s'-s, 0)$. Knowing $P(s')$, this implies that all these sets are pairwise disjoint. In addition, they cover $\mathscr{S}_{s'} \cap \{x \leq s-1\} - (s'-s, 0)$, which is none other that $\mathscr{S}_s$. Thus, $P(s)$ is proved. □

The key property of this construction is that the corresponding value of $c = n_0 \sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i) + m_s \sum_{i=1}^{s-1}(n_{i-h_i} - n_i)$ is softly linear in $n_0 m_s$. This is close to optimal, since the inequalities $\sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i) \geq m_s - 1$ and $\sum_{i=1}^{s-1}(n_{i-h_i} - n_i) \geq n_0 - 1$ imply that c is in $\Omega(n_0 m_s)$.

Proposition 4.4.2. *For* $\mathsf{R}_1, \ldots, \mathsf{R}_{s-1}$ *as above,* $c = n_0 \sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i) + m_s \sum_{i=1}^{s-1}(n_{i-h_i} - n_i)$ *is in* $O\tilde{}(n_0 m_s)$.

Proof. We prove that with the choices in Definition 27, $\sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i)$ is in $O\tilde{}(m_s)$; we omit the remaining part of the argument that proves that $\sum_{i=1}^{s-1}(n_{i-h_i} - n_i)$ is in $O\tilde{}(n_0)$ in a similar manner.

First, we reduce to the case where s is a power of 2. For $i \geq s$, set $\ell_i = 0$ and $m_i = m_s$; the sum $\sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i)$ is then equal to $\sum_{i=1}^{s'-1}(m_{i+\ell_i} - m_i)$, where $s' = 2^k$ is the first power of two greater than or equal to s. Besides, this convention implies $m_{i+\ell_i} = m_{i+h_i}$ for all i.

For a given κ in $\{0, \ldots, k-1\}$, the indices $i \in \{1, \ldots, s'-1\}$ of 2-adic valuation κ are the integers $2^\kappa(1+2j)$, for $j = 1, \ldots, 2^{k-\kappa-1} - 1$, so we can rewrite the sum $\sum_{i=1}^{s'-1}(m_{i+\ell_i} - m_i)$ as

$$\sum_{\kappa=0}^{k-1} \sum_{j=0}^{2^{k-\kappa-1}-1} (m_{2^\kappa(1+2j)+2^\kappa} - m_{2^\kappa(1+2j)}) = \sum_{\kappa=0}^{k-1} \sum_{j=0}^{2^{k-\kappa-1}-1} (d_{2^\kappa(1+2j)+1} + \cdots + d_{2^\kappa(1+2j)+2^\kappa}),$$

where we set $d_i = 0$ for $i > s$. In particular, for a fixed κ, the last index occurring at summation step j is less than the first index occurring at $j+1$, so the inner sum is bounded above by $\sum_{i=1}^{s'} d_i = m_s$. It follows that $\sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i) \le \sum_{\kappa=0}^{k-1} m_s \in O(m_s \log(s))$. Since $s \le m_s$, our claim is proved. □

4.4.2 Mixed radix representation

In this subsection, we discuss an alternative basis for our polynomials. Our motivation is the following: if $\mathcal{G} = (g_0, \dots, g_s)$ is the minimal, reduced lexicographic Gröbner basis that we want to use in our reduction algorithm, we saw that for $i = 0, \dots, s$, g_i can be written as $M_i G_i$, with M_i of degree m_i in $\mathbb{K}[x]$ and $G_i \in \mathbb{K}[x,y]$ monic in y, of degree n_i in y. Recall also that for $i = 1, \dots, s$ we write $D_i = M_i / M_{i-1}$, which is a polynomial of degree $d_i = m_i - m_{i-1}$ in $\mathbb{K}[x]$.

The main reduction algorithm will perform many univariate reductions modulo the polynomials $M_1, \dots, M_s$. When working with $\langle x, y\rangle$-primary ideals, all M_i's are powers of x, so these operations are free of arithmetic cost. In general, though, this is not the case anymore, if the inputs are represented on the monomial basis. In this paragraph, we introduce a mixed radix representation where reductions by the M_i's are free, and we discuss conversion algorithms.

Given polynomial $\boldsymbol{K} = (K_1, \dots, K_t)$ in $\mathbb{K}[x]$, with respective degrees $k_1, \dots, k_t$, and writing $h = k_1 + \cdots + k_t$, we consider the $\mathbb{K}$-linear mapping

$$\begin{aligned} \Phi_{\boldsymbol{K}} : \mathbb{K}[x]_{<k_1} \times \cdots \times \mathbb{K}[x]_{<k_t} &\to \mathbb{K}[x]_{<h} \\ (F_1, \dots, F_t) &\mapsto F_1 + K_1 F_2 + K_1 K_2 F_3 + \cdots + K_1 \cdots K_{t-1} F_t. \end{aligned}$$

The domain and codomain both have dimension h; from this, we easily deduce that $\Phi_{\boldsymbol{K}}$ is a $\mathbb{K}$-vector space isomorphism. For F in $\mathbb{K}[x]_{<h}$, we call $(F_1, \dots, F_t) = \Phi_{\boldsymbol{K}}^{-1}(F)$ its *mixed radix* representation with respect to the basis $\boldsymbol{K}$.

We will rely on the following fact: given $(F_1, \dots, F_t) = \Phi_{\boldsymbol{K}}^{-1}(F)$, for i in $\{1, \dots, t\}$, the mixed radix representation of F div $K_1 \cdots K_i$, with respect to the basis $(K_{i+1}, \dots, K_t)$, is $(F_{i+1}, \dots, F_t)$, so we have access to it free of cost. Similarly, the mixed radix representation of F rem $K_1 \cdots K_i$, with respect to the basis $(K_1, \dots, K_i)$, is $(F_1, \dots, F_i)$. In particular, if F is given in its mixed radix representation, quotient and remainder by the product $K_1 \cdots K_i$ are free; we still denote these operations by div and rem.

Conversely, for F of degree less than $k_{i+1} + \cdots + k_t$, given on the mixed radix basis associated to $(K_{i+1}, \dots, K_t)$ as a vector $(F_{i+1}, \dots, F_t)$, the mixed radix representation of $K_1 \cdots K_i F$, for the basis $(K_1, \dots, K_t)$, is $(0, \dots, 0, F_{i+1}, \dots, F_t)$, so it can be computed for free.

For completeness, we give algorithms with softly linear runtime to apply $\Phi_{\boldsymbol{K}}$ and its inverse. These are elementary variants of the algorithms for Chinese remaindering in [73, Chapter 10.3], or generalized Taylor expansion [73, Chapter 9.2]. We start with the conversion from the mixed radix to monomial representation.

Algorithm 4.4.1 FromMixedRadix$((F_1, \ldots, F_t), (K_1, \ldots, K_t))$

```
Input: $(F_1, \ldots, F_t)$ in $\mathbb{K}[x]_{<k_1} \times \cdots \times \mathbb{K}[x]_{<k_t}$, $\boldsymbol{K} = (K_1, \ldots, K_t)$ of respective degrees
    $k_1, \ldots, k_t$
Output: $\Phi_{\boldsymbol{K}}(F_1, \ldots, F_t) \in \mathbb{K}[x]_{<h}$, with $h = k_1 + \cdots + k_t$
1: if $t = 1$ then return $F_1$
2: $t' \leftarrow \lceil t/2 \rceil$
3: $L \leftarrow$ FromMixedRadix$((F_1, \ldots, F_{t'}), (K_1, \ldots, K_{t'}))$
4: $R \leftarrow$ FromMixedRadix$((F_{t'+1}, \ldots, F_t), (K_{t'+1}, \ldots, K_t))$
5: if $R = 0$ then
6:     return $L$
7: else
8:     return $L + K_1 \cdots K_{t'} R$
```

Correctness is clear: if we write $F = \Phi_{\boldsymbol{K}}(F_1, \ldots, F_t)$, then the previous discussion shows that $L = F \text{ rem } K_1 \cdots K_{t'}$ and $R = F \text{ div } K_1 \cdots K_{t'}$, so that the output is indeed F. If we enter Line 8, computing P takes $O^\sim(k_1 + \cdots + k_{t'})$ operations $(+, \times)$ in $\mathbb{K}$ [73, Lemma 10.4]; however, in this case R is nonzero, so F has degree at least $k_1 + \cdots + k_{t'}$, and $O^\sim(k_1 + \cdots + k_{t'})$ is $O^\sim(\deg(F))$. It follows that, excluding the recursive calls, the cost of a single call to Algorithm FromMixedRadix is $O^\sim(\deg(F))$ if $\deg(F) \geq k_1 + \cdots + k_{t'}$, and zero otherwise.

There are $O(\log(\deg(F)))$ levels of the recursion tree that will incur a nonzero cost, and the degrees of the polynomials computed at any of these levels add up to at most $\deg(F)$. Hence, the overall cost is $O^\sim(\deg(F))$ operations $(+, \times)$ in $\mathbb{K}$.

For the inverse operation, the algorithm is recursive as well. Using the test at Line 3, we avoid doing any computation if F has degree less than $k_1 + \cdots + k_{t'}$. The discussion is as above, yielding a runtime of $O^\sim(\deg(F))$ operations $(+, \times)$ in $\mathbb{K}$.

Algorithm 4.4.2 ToMixedRadix$(F, (K_1, \dots, K_t))$

```
INPUT: $F$ in $\mathbb{K}[x]_{<h}$, $\boldsymbol{K} = (K_1, \dots, K_t)$ of respective degrees $k_1, \dots, k_t$, with $h = k_1 + \cdots + k_t$
OUTPUT: $(F_1, \dots, F_t) = \Phi_{\boldsymbol{K}}^{-1}(F)$
if $t = 1$ then return $(F)$
$t' \leftarrow \lceil t/2 \rceil$
if $\deg(F) < k_1 + \cdots + k_{t'}$ then
    return ToMixedRadix$(F, (K_1, \dots, K_{t'}))$ cat $(0, \dots, 0)$        ▷ $t - t'$ zeros
else
    $P \leftarrow K_1 \cdots K_{t'}$
    $Q, R \leftarrow F$ div $P, F$ rem $P$
    return ToMixedRadix$(R, (K_1, \dots, K_{t'}))$ cat ToMixedRadix$(Q, (K_{t'+1}, \dots, K_t))$
```

In the next paragraphs, we apply these algorithms to polynomials in $\mathbb{K}[x, y]$ (we use the same names for the algorithms). In this case, we simply proceed coefficient-wise with respect to y, the mixed-radix representation of $F \in \mathbb{K}[x, y]$ being now a two-dimensional array. If the sum of the degrees of $K_1, \dots, K_t$ is h, and for F in $\mathbb{K}[x, y]$ supported on an initial segment U, with also $\deg(F, x) < h$, the runtime of both algorithms is $O\tilde{~}(|\mathsf{U}|)$.

4.4.3 The main algorithm

We can now use the results from the previous subsections in order to give an algorithm for the reduction of a polynomial $f \in \mathbb{K}[x, y]$ modulo a minimal reduced lexicographic Gröbner basis $\mathcal{G} = (g_0, \dots, g_s)$. For the time being, we only consider the "balanced" case, where f is already reduced modulo g_0 and g_s. Let us write, as usual, the initial terms of $\mathcal{G}$ as

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \dots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$$

with the m_i's increasing and the n_i's decreasing, and let S be the rectangle $\{0, \dots, m_s - 1\} \times \{0, \dots, n_0 - 1\}$. Then, our assumption is that f is in $\mathbb{K}[x, y]_{\mathsf{S}}$. More general inputs can be handled by performing a reduction by (g_0, g_s) first; this is discussed in the last paragraph of this section.

In what follows, we let T be the initial segment determined by $\mathcal{G}$, and $\delta = \dim_{\mathbb{K}}(\mathbb{K}[x, y]/\mathcal{G})$ be the degree of $\mathcal{G}$.

4.4.3.1 Overview of the algorithm Given f in $\mathbb{K}[x, y]$ with $\deg(f, x) < m_s$ and $\deg(f, y) < n_0$, our main algorithm Reduction computes $r = f$ rem $\mathcal{G}$ by calling

$s-1$ times a procedure called PartialReduction, which is described further. The main algorithm returns the remainder r, together with quotients $Q_1, \dots, Q_{s-1}$, such that $f = Q_1 g_1 + \cdots + Q_s g_s + r$. While we do not need the quotients in this paper, we return them as a byproduct that could possibly be of use in other contexts (the algorithm does not compute the last quotient Q_s, but it would be straightforward to deduce it from the output, if needed). Since we assume $\deg(f, y) < n_0$, g_0 does not appear in the reduction equality.

The mixed radix basis is used throughout the algorithm to handle intermediate data; input and output are on the usual monomial basis.

Algorithm 4.4.3 Reduction$(f, \mathcal{G})$

```
Input: $f$ in $\mathbb{K}[x, y]$, $\mathcal{G} = (g_0, \dots, g_s)$ as above
Assumptions: $\deg(f, x) < m_s$, $\deg(f, y) < n_0$
Output: $f$ rem $\mathcal{G}$ and quotients $Q_1, \dots, Q_{s-1}$
$M_0 \leftarrow 1$, $G_0 \leftarrow g_0$
for $i = 1, \dots, s$ do
    $M_i \leftarrow$ PolynomialCoefficient$(g_i, y^{n_i}) \in \mathbb{K}[x]$
    $G_i \leftarrow g_i$ div $M_i$
    $D_i \leftarrow M_i$ div $M_{i-1}$
$f^{(0)} \leftarrow$ ToMixedRadix$(f, (D_1, \dots, D_s))$            ▷ $f^{(0)}$ is on the mixed radix basis
for $i = 1, \dots, s-1$ do
    $f^{(i)}, Q_i \leftarrow$ PartialReduction$(f^{(i-1)}, i)$ ▷ all $f^{(i)}$ are on the mixed radix basis
return FromMixedRadix$(f^{(s-1)}, D_1, \dots, D_s), Q_1, \dots, Q_{s-1}$
```

To simplify notation, the polynomials $g_0, \dots, g_s$, $G_0, \dots, G_s$, $M_0, \dots, M_s$ and $D_1, \dots, D_s$, the latter of which are computed at the beginning of the main algorithm, are assumed to be known in our calls to Algorithm PartialReduction, rather than passed as arguments.

The main result in this section is the following proposition. The runtime given here is softly linear in $n_0 m_s$ and $s\delta$: the former represents the size of the input polynomial f, and the latter is the upper bound on the number of coefficients needed to represent $\mathcal{G}$ discussed in Section 4.3.1. Whether a better algorithm is possible (which would not need all coefficients of $\mathcal{G}$, but only, for instance, its Gröbner parameters) is not clear to us.

Proposition 4.4.3. *Given f and $\mathcal{G}$, with $\deg(f, x) < m_s$ and $\deg(f, y) < n_0$, Algorithm* Reduction *returns f rem $\mathcal{G}$ using $O\tilde{~}(n_0 m_s + s\delta)$ operations $(+, \times)$ in $\mathbb{K}$.*

Before proving the proposition, we mention an important particular case, where a simplified runtime is available. Suppose that $e_i = 1$ for all i, that is, that all steps in the staircase have height 1. In this case, $n_0 = s$, and since we have $m_s \le \delta$, we obtain $n_0 m_s \le s\delta$. In other words, the runtime of the algorithm is simply $O\tilde{~}(s\delta)$.

4.4.3.2 A single reduction step

We start with a description of the key subroutine, Algorithm PartialReduction.

In Section 4.4.1, we described a way to cover $\mathsf{S} = \{0, \dots, m_s - 1\} \times \{0, \dots, n_0 - 1\} - \mathsf{T}$ by rectangles $\mathsf{R}_i = \{m_i, \dots, m_{i+\ell_i} - 1\} \times \{n_i, \dots, n_{i-h_i} - 1\}$, with $h_i = 2^{\mathrm{val}_2(i)}$ and $\ell_i = \min(h_i, s - i)$ for $i = 1, \dots, s - 1$.

In Algorithm PartialReduction, we are given $f \in \mathbb{K}[x, y]_{\mathsf{S}}$, and an index i in $\{1, \dots, s - 1\}$. The essential operation is a Euclidean division with respect to the variable y, with coefficients suitably reduced with respect to x; the mixed radix basis is used to control the cost of this reduction in x. We prove below that the output r has the same remainder as f modulo $\mathcal{G}$; we also return the partial quotient Q, which is supported on a translate of R_i.

Lemma 4.4.1. *Calling* PartialReduction(f, i) *takes*

$$O\tilde{~}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$$

operations $(+, \times)$ *in* $\mathbb{K}$, *with* $h_i = 2^{\mathrm{val}_2(i)}$ *and* $\ell_i = \min(h_i, s - i)$. *The output* r, Q *satisfies the following properties:*

1. $\deg(r, x) < m_s$ *and* $\deg(r, y) < n_0$
2. $r \operatorname{rem} \mathcal{G} = f \operatorname{rem} \mathcal{G}$
3. $r \operatorname{rem} M_i = f \operatorname{rem} M_i$
4. $r \operatorname{div} y^{n_{i-h_i}} = f \operatorname{div} y^{n_{i-h_i}}$
5. $((r \operatorname{div} M_i) \operatorname{div} y^{n_i}) \operatorname{rem} (D_{i+1} \cdots D_{i+\ell_i}, y^{n_{i-h_i} - n_i}) = 0$
6. $\deg(Q, x) < m_{i+\ell_i} - m_i$ *and* $\deg(Q, y) < n_{i-h_i} - n_i$.

Proof. We first verify that all steps are well-defined, and discuss degree properties of the polynomials in the algorithm.

As per our discussion in the preamble, the division and remainder at Lines 2 and 3 output a bivariate polynomial F_2 on the mixed radix basis associated to $D_{i+1}, \dots, D_{i+\ell_i}$. The polynomial F_3 is written on the same basis; F_4 represents the same polynomial, this time on the monomial basis.

Algorithm 4.4.4 PartialReduction(f, i)

```
Input: $f$ in $\mathbb{K}[x, y]$, $i$ in $\{1, \ldots, s-1\}$
Assumptions: $\deg(f, x) < m_s$, $\deg(f, y) < n_0$, $i$ in $\{1, \ldots, s-1\}$. $f$ is given on the
    mixed radix basis associated to $D_1, \ldots, D_s$
Output: $r$ and $Q$ in $\mathbb{K}[x, y]$. $r$ is given on the mixed radix basis associated to
    $D_1, \ldots, D_s$
$h_i \leftarrow 2^{\mathrm{val}_2(i)}$, $\ell_i \leftarrow \min(h_i, s-i)$
$F_1 \leftarrow f$ div $M_i$    ▷ division in the mixed radix basis
    ▷ $F_1$ is given on the mixed radix basis associated to $D_{i+1}, \ldots, D_s$
$F_2 \leftarrow F_1$ rem $D_{i+1} \cdots D_{i+\ell_i}$    ▷ division in the mixed radix basis
    ▷ $F_2$ is given on the mixed radix basis associated to $D_{i+1}, \ldots, D_{i+\ell_i}$
$F_3 \leftarrow F_2$ rem $y^{n_{i-h_i}}$
$F_4 \leftarrow$ FromMixedRadix$(F_3, (D_{i+1}, \ldots, D_{i+\ell_i}))$    ▷ $F_4$ is on the monomial basis
$q \leftarrow F_4$ div $G_i$ in $\mathbb{A}[y]$    ▷ $G_i$ such that $g_i = M_i G_i$, $\mathbb{A} = \mathbb{K}[x]/\langle D_{i+1} \cdots D_{i+\ell_i} \rangle$
let $Q$ be the canonical lift of $q$ to $\mathbb{K}[x, y]$    ▷ $\deg(Q, x) < m_{i+\ell_i} - m_i$
$V \leftarrow$ Multiply$(Q, \{0, \ldots, m_{i+\ell_i} - m_i - 1\} \times \{0, \ldots, n_{i-h_i} - n_i - 1\}, G_i, \mathsf{T})$
    ▷ $V = QG_i$ on the monomial basis
$V_1 \leftarrow V$ rem $(D_{i+1} \cdots D_s)$    ▷ $V_1 = QG_i$ rem $(D_{i+1} \cdots D_s)$ on the monomial basis
$V_2 \leftarrow$ ToMixedRadix$(V_1, (D_{i+1}, \ldots, D_s))$
    ▷ $V_2 = QG_i$ rem $(D_{i+1} \cdots D_s)$, given on the mixed radix basis associated to
    $D_{i+1}, \ldots, D_s$
$V_3 \leftarrow M_i V_2$    ▷ multiplication in the mixed radix basis
    ▷ $V_3 = Qg_i$ rem $M_s$, given on the mixed radix basis associated to $D_1, \ldots, D_s$
$r \leftarrow f - V_3$    ▷ subtraction in the mixed radix basis
    ▷ $r = (f - Qg_i)$ rem $M_s$, given on the mixed radix basis associated to $D_1, \ldots, D_s$
return $r, Q$
```

That polynomial has y-degree less than n_{i-h_i}; since G_i has y-degree n_i, q, and thus Q, have y-degree less than $n_{i-h_i} - n_i$. Since Q also has x-degree less than $m_{i+\ell_i} - m_i$, it is supported on the rectangle $\{0, \dots, m_{i+\ell_i} - m_i - 1\} \times \{0, \dots, n_{i-h_i} - n_i - 1\}$ (which is the translate of R_i to the origin). This proves the last claim in the lemma.

On the other hand, G_i is supported on T (this is true because $i \geq 1$; for $i = 0$, the initial term of G_0, which is y^{n_0}, is not in T), so altogether, the call to Multiply at Line 8 is justified. The variables V_1 and V_2 then represent the same polynomial, namely QG_i rem $(D_{i+1} \cdots D_s)$, on two different bases (resp. monomial and mixed radix). It follows that V_3 represents the polynomial

$$\begin{aligned} M_i(QG_i \text{ rem } (D_{i+1} \cdots D_s)) &= M_iQG_i \text{ rem } (M_iD_{i+1} \cdots D_s) \\ &= Qg_i \text{ rem } M_s. \end{aligned}$$

As we noted in the previous subsection, since V_2 is written on the mixed basis associated to $(D_{i+1}, \dots, D_s)$, V_3 is written on the mixed basis associated to $(D_1, \dots, D_s)$. Since this is also the case for f, the subtraction at Line 12 is done coefficient-wise, and results in the polynomial $(f - Qg_i)$ rem M_s, written on the same mixed basis.

This being said, we establish properties 1-5. First item: We have $\deg(f, y) < n_0$. On the other hand, the degree bound on Q implies that Qg_i has y-degree less than n_{i-h_i}. Since $n_{i-h_i} \leq n_0$, the product Qg_i has y-degree less than n_0 as well, and it is then also the case for r. The bound $\deg(r, x) < m_s$ holds by construction.

Second item: we can write $r = f - Qg_i + hM_s = f - Qg_i + hg_s$, for some h in $\mathbb{K}[x, y]$, so that $r - f$ is in the ideal $\langle \mathcal{G} \rangle$.

Third item: consider again the expression $r = f - Qg_i + hM_s$, and notice that M_i divides both g_i and M_s.

Fourth item: because $\deg(f, x) < m_s$, the quotient h in the relation $r = f - Qg_i + hM_s$ is $-Qg_i$ div M_s. Since Qg_i has y-degree less than n_{i-h_i}, it is thus also the case for h. This shows that r div $y^{n_{i-h_i}} = f$ div $y^{n_{i-h_i}}$, as claimed.

Fifth item: since $r = f - Qg_i + hM_s = f - QM_iG_i + hM_s$, we have r div $M_i = F_1 - QG_i + hD_{i+1} \cdots D_s$. By definition, we have $F_1 = F_2 + LD_{i+1} \cdots D_{i+\ell_i}$ and $F_2 = F_3 + Ky^{n_{i-h_i}}$ for some K, L in $\mathbb{K}[x, y]$. F_4 is the same polynomial as F_3, written on a different basis, and satisfies $F_4 = QG_i + P + L'D_{i+1} \cdots D_{i+\ell_i}$, for some P and L' in $\mathbb{K}[x, y]$, with P of y-degree less than n_i. Altogether, we obtain r div $M_i = P + (L + L')D_{i+1} \cdots D_{i+\ell_i} + hD_{i+1} \cdots D_s + Ky^{n_{i-h_i}}$. As a result,

$$(r \text{ div } M_i) \text{ div } y^{n_i} = ((L+L') \text{ div } y^{n_i})D_{i+1} \cdots D_{i+\ell_i} + (h \text{ div } y^{n_i})D_{i+1} \cdots D_s + Ky^{n_{i-h_i} - n_i}.$$

Because $i + \ell_i \leq s$, this expression taken modulo $(D_{i+1} \cdots D_{i+\ell_i}, y^{n_{i-h_i} - n_i})$ vanishes, as claimed.

It remains to estimate the cost of the algorithm. The divisions with remainders at Lines 2 and 3 are free of cost (because we work in the suitable mixed radix bases); the same holds for Line 4, since it only involves a power of y.

Since $D_{i+1} \cdots D_{i+\ell_i}$ has degree $m_{i+\ell_i} - m_i$, the conversion at Line 5 uses $O\tilde{}(n_{i-h_i}(m_{i+\ell_i} - m_i))$ operations $(+, \times)$ in $\mathbb{K}$, which is $O\tilde{}(n_0(m_{i+\ell_i} - m_i))$.

Prior to the division at Line 6, G_i has to be reduced modulo $D_{i+1} \cdots D_{i+\ell_i}$; proceeding coefficient-wise in y, this takes $O\tilde{}(|\mathsf{T}|) = O\tilde{}(\delta)$ operations $(+, \times)$ in $\mathbb{K}$. Then, the division in $\mathbb{A}[y]$ takes $O\tilde{}(n_{i-h_i})$ operations $(+, \times)$ in $\mathbb{A}$, which is $O\tilde{}(n_{i-h_i}(m_{i+\ell_i} - m_i))$ operations $(+, \times)$ in $\mathbb{K}$. For this expression, it will be enough to use the same upper bound $O\tilde{}(n_0(m_{i+\ell_i} - m_i))$ as above.

Next, we consider the cost of computing the product V in $\mathbb{K}[x, y]$. The input Q has x-degree less than $m_{i+\ell_i} - m_i$ and y-degree less than $n_{i-h_i} - n_i$, whereas G_i is supported on the initial segment T of height n_0, width m_s, and cardinal δ. Hence, using Proposition 4.2.2 (and the remarks that follow the proposition on the size of the support of QG_i), we see that QG_i can be computed in $O\tilde{}((m_{i+\ell_i} - m_i)(n_{i-h_i} - n_i) + n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$ operations $(+, \times)$ in $\mathbb{K}$. This is also $O\tilde{}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$.

The Euclidean division at Line 9 is done in the monomial basis, proceeding coefficient-wise in y. Computing $D_{i+1} \cdots D_s$ takes $O\tilde{}(m_s)$ operations $(+, \times)$ in $\mathbb{K}$. Then, the reduction is done in quasi-linear time in the size of the support of V, that is, $O\tilde{}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$ again. Recall that for polynomials supported on an initial segment U, the conversion to the mixed radix basis takes quasi-linear time in the size of U. Here, the support U is contained in the support of $V = QG_i$, so the conversion at Line 10 takes time $O\tilde{}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$ once more.

The multiplication by M_i in the mixed radix basis is free, as we simply prepend a vector of zeros to each entry of V_2 to obtain V_3. Finally, the polynomial subtraction at the last step involves one subtraction in $\mathbb{K}$ for each nonzero coefficient of V_3, so $O\tilde{}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$ altogether. □

4.4.3.3 Correctness of the main algorithm

The properties stated above allow us to prove that Algorithm Reduction correctly computes the remainder of f by $\mathcal{G}$.

We define indices $(b_{i,j})_{0 \le i < s, 0 \le j < n_0}$ in $\{1, \ldots, s\}$ as follows. For $i = 0, \ldots, s-1$, let $\mathsf{T}_i \subset \mathbb{N}^2$ be the union of the initial segment T and the rectangles $\mathsf{R}_1, \ldots, \mathsf{R}_i$; in particular, $\mathsf{T}_0 = \mathsf{T}$ and T_{s-1} is the rectangle $\{0, \ldots, m_s - 1\} \times \{0, \ldots, n_0 - 1\}$. Then, for $i = 0, \ldots, s-1$ and $j = 0, \ldots, n_0 - 1$, we let $b_{i,j} \in \{1, \ldots, s\}$ be the smallest index k such that (m_k, j) is not in T_i. In particular, $b_{s-1,j} = s$ for all $j < n_0$. On the other hand, for $i = 0$, we see that any pair (u, j) with $u < m_{b_{0,j}}$ is in T, so $x^u y^j$ is reduced modulo $\mathcal{G}$.

Let $f^{(0)}, \ldots, f^{(s-1)}$ be the polynomials computed throughout the algorithm (the first item of Lemma 4.4.1 proves that these polynomials are well-defined, and all supported on the rectangle $\{0, \ldots, m_s - 1\} \times \{0, \ldots, n_0 - 1\}$). We prove the following claim, written $A(i)$ in the sequel, by induction on $i = 0, \ldots, s-1$: for $n_i \le j < n_0$, the polynomial $\textsc{PolynomialCoefficient}(f^{(i)}, y^j) \text{ rem } M_{b_{i,j}} \in \mathbb{K}[x]$ has degree less than $m_{b_{0,j}}$. For $i = 0$, there is nothing to prove (since no index j needs to be considered). Suppose that $A(i-1)$ holds, for some i in $\{1, \ldots, s-1\}$; we prove $A(i)$.

For $j \ge n_{i-h_i}$, Item 4 of Lemma 4.4.1 shows that $\textsc{PolynomialCoefficient}(f^{(i)}, y^j) = \textsc{PolynomialCoefficient}(f^{(i-1)}, y^j)$. Since in that case we also have $b_{i,j} = b_{i-1,j}$, our claim holds. Now, suppose that j is in $\{n_i, \ldots, n_{i-h_i} - 1\}$; in this case, Items 3 and 5 of the same lemma imply that $\textsc{PolynomialCoefficient}(f^{(i)}, y^j) \text{ rem } M_{i+\ell_i}$ is equal to $\textsc{PolynomialCoefficient}(f^{(i-1)}, y^j) \text{ rem } M_i$. On the other hand, we also have $b_{i-1,j} = i$ and $b_{i,j} = i + \ell_i$, so the left-hand side is the term $\textsc{PolynomialCoefficient}(f^{(i)}, y^j) \text{ rem } M_{b_{i,j}}$ that appears in our claim. Thus, to conclude the induction proof, it is enough to show that $\textsc{PolynomialCoefficient}(f^{(i-1)}, y^j) \text{ rem } M_i$ has degree less than $m_{b_{0,j}}$. We do this using a further case discussion:

- if $j \ge n_{i-1}$, we can use the induction assumption. It implies that the remainder $\textsc{PolynomialCoefficient}(f^{(i-1)}, y^j) \text{ rem } M_{b_{i-1,j}}$ has degree less than $m_{b_{0,j}}$. Since we saw that have $b_{i-1,j} = i$, we are done.

- if $j < n_{i-1}$, we have $b_{0,j} = i$, so that $m_{b_{0,j}} = m_i = \deg(M_i)$, and our claim holds as well.

Having established our induction claim, we can take $i = s-1$. Then, $A(s-1)$ shows that for j in $n_{s-1}, \ldots, n_0 - 1$, $\textsc{PolynomialCoefficient}(f^{(s-1)}, y^j) \text{ rem } M_s$ has degree less than $m_{b_{0,j}}$. By construction, $f^{(s-1)}$ is reduced modulo M_s, so that $\textsc{PolynomialCoefficient}(f^{(s-1)}, y^j)$ itself has degree less than $m_{b_{0,j}}$. Now, for j in $0, \ldots, n_{s-1} - 1$, we have $b_{0,j} = s$, so $\textsc{PolynomialCoefficient}(f^{(s-1)}, y^j)$ has degree less than $m_{b_{0,j}}$ as well in this case. Altogether, as we pointed out when we introduced $m_{b_{0,j}}$, this proves that $f^{(s-1)}$ is reduced modulo $\mathcal{G}$.

The second item of Lemma 4.4.1 finally shows that $f \text{ rem } \mathcal{G} = f^{(s-1)} \text{ rem } \mathcal{G}$, so $f^{(s-1)}$ is indeed the normal form of f modulo $\mathcal{G}$. This finishes the correctness proof.

4.4.3.4 Cost analysis For the cost analysis, we start with the computation of polynomials M_i, G_i and D_i, at the beginning of the main algorithm. Since division by a monic univariate polynomial take softly linear time, each pass in the loop at Line 2 of Reduction takes $\tilde{O}(\delta)$ operations, for a total of $\tilde{O}(s\delta)$.

The conversions to and from the mixed radix basis take quasi-linear time in the size of the support of f, that is, $\tilde{O}(n_0 m_s)$ operations. Then, it suffices to add the costs of the calls to PartialReduction. By Lemma 4.4.1, deducing $f^{(i)}$ from $f^{(i-1)}$ takes $\tilde{O}(n_0(m_{i+\ell_i} - m_i) + m_s(n_{i-h_i} - n_i) + \delta)$ operations in $\mathbb{K}$, with $\delta = |\mathsf{T}|$, so it suffices to sum this quantity for $i = 1$ to $s-1$. The first two terms add up to a total of $\tilde{O}(n_0 \sum_{i=1}^{s-1}(m_{i+\ell_i} - m_i) + m_s \sum_{i=1}^{s-1}(n_{i-h_i} - n_i))$. Proposition 4.4.2 shows that this sum is in $\tilde{O}(n_0 m_s)$, so taking into account the term $\tilde{O}(\delta)$ in each summand, the total is $\tilde{O}(n_0 m_s + s\delta)$, as claimed.

4.4.3.5 Generalization to arbitrary inputs and discussion If the input f does not satisfy the conditions $\deg(f, x) < m_s$ and $\deg(f, y) < n_0$, we fall back to this case by reduction modulo the pair of polynomials (g_0, g_s), which have respective initial terms y^{n_0} and x^{m_s}. The following straightforward algorithm achieves this; we discuss possible improvements below.

Algorithm 4.4.5 ReductionGeneralInput$(f, \mathcal{G})$

```
Input: $f$ in $\mathbb{K}[x,y]$, $\mathcal{G} = (g_0, \dots, g_s)$
Output: $f$ rem $\mathcal{G}$
$f_1 \leftarrow f$ rem $g_s$
$f_2 \leftarrow f$ rem $g_0$ in $\mathbb{A}[y]$                        ▷ $\mathbb{A} = \mathbb{K}[x]/\langle g_s \rangle$
let $f_3$ be the canonical lift of $f_2$ to $\mathbb{K}[x,y]$      ▷ $\deg(f_3, x) < m_s$
return Reduction$(f_3, \mathcal{G})$
```

Proposition 4.4.4. *Given f and $\mathcal{G}$, with $\deg(f, x) < d$ and $\deg(f, y) < e$, Algorithm ReductionGeneralInput returns f rem $\mathcal{G}$ using $\tilde{O}(ed + em_s + n_0 m_s + s\delta)$ operations $(+, \times)$ in $\mathbb{K}$. If $\mathcal{G}$ generates an $\langle x, y\rangle$-primary ideal, the runtime becomes $\tilde{O}(\delta m_s)$ operations $(+, \times)$ in $\mathbb{K}$.*

Proof. Reducing f modulo g_s takes $\tilde{O}(ed)$ operations (and is actually free if $d < m_s$). Then, Euclidean division by g_0 in $\mathbb{A}[y]$ uses $\tilde{O}(e)$ steps in $\mathbb{A}$, which is $\tilde{O}(em_s)$ steps in $\mathbb{K}$. Finally, Proposition 4.4.3 gives a cost of $\tilde{O}(n_0 m_s + s\delta)$ for the last step.

If $\mathcal{G}$ generates an $\langle x, y\rangle$-primary ideal, all terms of y-degree at least δ vanish through the reduction (so we can replace e by δ), as do all terms of x-degree at least m_s (so we can replace d by m_s). □

In the runtime for the general case, ed is the size of the support of input f, and $s\delta$ our bound on the size of $\mathcal{G}$, so they are essentially unavoidable (unless of course one could avoid using $\mathcal{G}$ itself but only its Gröbner parameters). The runtime also features the extra terms em_s and n_0m_s, but getting rid of them and improving the runtime to $O\tilde{}(ed + s\delta)$ unconditionally seems to be very challenging.

Indeed, consider the *modular composition* problem: given F, G, H in $\mathbb{K}[x]$, with F monic of degree n and G, H of degrees less than n, this amounts to computing $G(H)$ rem F. A direct approach takes quadratic time, and Brent-Kung's baby-steps / giant-steps algorithm uses $O(n^{1.69})$ operations (and relies on fast matrix arithmetic). Bringing this down to a quasi-linear runtime has been an open question since 1978: it is so far known to be feasible only over finite $\mathbb{K}$ [101], with the best algorithm for an arbitrary $\mathbb{K}$ to date featuring a Las Vegas cost of $O(n^{1.43})$ [129].

It turns out that modular composition is a particular case of the reduction problem we are considering here. With F, G, H as above, if we consider $\mathcal{G} = (y - H(x), F(x))$ and the polynomial $f = G(y)$, then the remainder f rem $\mathcal{G}$ is precisely $G(H)$ rem F. Here, we have $n_0 = 1$, $s = 1$, $m_s = n$, $\delta = n$, $d = 1$ and $e = \deg(G, y) + 1$, so that in general $e = n$; on such input, the runtime of our algorithm is $O\tilde{}(n^2)$. Improving our result to $O\tilde{}(ed + s\delta)$ would give a softly linear modular composition algorithm, thus solving a long-standing open question.

On the other hand, the case where f has a large degree in both x and y, *i.e.* when $m_s \leq d$ and $n_0 \leq e$, is particularly favourable, since then the runtime does become $O\tilde{}(ed + s\delta)$. Another favourable situation is when all e_i's are equal to 1, since we said before that we have $n_0m_s \leq s\delta$ in this case, with thus a runtime of $O\tilde{}(ed + em_s + s\delta)$.

Finally, we point out an application of Proposition 4.4.4 to modular multiplication: given A, B in $\mathbb{K}[x, y]_{\mathsf{T}}$, where T is the initial segment determined by $\mathcal{G}$, compute $f = AB$ rem $\mathcal{G} \in \mathbb{K}[x, y]_{\mathsf{T}}$. In this case, we have $d < 2m_s$ and $e < 2n_0$, so the runtime is $O\tilde{}(n_0m_s + s\delta)$; when all e_i's are equal to 1, this becomes $O\tilde{}(s\delta)$. We are not aware of previous results for this question.

4.5 From Gröbner parameters to Gröbner basis

In this section, we fix a given Gröbner cell (or equivalently, the monomials $\boldsymbol{E}$). We show how make explicit the mapping $\Phi_{\boldsymbol{E}} : \mathbb{K}^N \to \mathcal{C}(\boldsymbol{E})$, which takes as input Gröbner parameters and outputs the corresponding reduced Gröbner basis (see Section 4.3.2).

First, we fix a way to index the N coefficients of the polynomials $(\sigma_{i,j})_{0 \leq i \leq s-1, i \leq j \leq s}$ that appear in the syzygy (4.3.0.2); this will be done in the mutually inverse routines given below. Here, for simplicity, we assume that given the monomials $\boldsymbol{E}$, we can directly access the integers s, $(d_i)_{1 \leq i \leq s}$ and $(e_i)_{1 \leq i \leq s}$.

Algorithm 4.5.1 SigmaFromParameters($\boldsymbol{E}, (\lambda_1, \ldots, \lambda_N)$)

```
Input: monomials $\boldsymbol{E}$, $(\lambda_1, \ldots, \lambda_N)$ in $\mathbb{K}^N$
Output: polynomials $(\sigma_{i,j})_{0 \leq i \leq s-1, i \leq js}$ in $\mathbb{K}[x, y]$
$k \leftarrow 1$
for $i = 0, \ldots, s-1$ do
    $\sigma_{i,i} \leftarrow \sum_{0 \leq \ell < d_{i+1}} \lambda_{k+\ell} x^\ell$
    $k \leftarrow k + d_{i+1}$
    for $j = i+1, \ldots, s$ do
        $\sigma_{i,j} \leftarrow 0$
        for $m = 0, \ldots, e_{j-1}$ do
            $\sigma_{i,j} \leftarrow \sigma_{i,j} + \sum_{0 \leq \ell < d_{i+1}} \lambda_{k+\ell} x^\ell y^m$
            $k \leftarrow k + d_{i+1}$
return $(\sigma_{i,j})_{0 \leq i \leq s-1, i \leq j \leq s}$
```

Algorithm 4.5.2 ParametersFromSigma($\boldsymbol{E}, (\sigma_{i,j})_{i,j}$)

```
Input: monomials $\boldsymbol{E}$, polynomials $(\sigma_{i,j})_{i,j}$ in $\mathbb{K}[x, y]$
Output: $(\lambda_1, \ldots, \lambda_N)$ in $\mathbb{K}^N$
$k \leftarrow 1$
for $i = 0, \ldots, s-1$ do
    for $\ell = 0, \ldots, d_{i+1} - 1$ do $\lambda_{k+\ell} \leftarrow$ Coefficient($\sigma_{i,i}, x^\ell$)
    $k \leftarrow k + d_{i+1}$
    for $j = i+1, \ldots, s$ do
        $\sigma_{i,j} \leftarrow 0$
        for $m = 0, \ldots, e_{j-1}$ do
            for $\ell = 0, \ldots, d_{i+1} - 1$ do $\lambda_{k+\ell} \leftarrow$ Coefficient($\sigma_{i,j}, x^\ell y^m$)
            $k \leftarrow k + d_{i+1}$
return $(\lambda_1, \ldots, \lambda_N)$
```

To deal with the particular case of punctual Gröbner parameters, a few obvious modifications are needed, such as setting $\sigma_{0,0}, \ldots, \sigma_{s-1,s-1}$ to zero and ensuring that x divides $\sigma_{1,0}, \ldots, \sigma_{s,s-1}$ in SigmaFromParameters. We call SigmaFromPunctualParameters and PunctualParametersFromSigma the resulting procedures.

We can now give an algorithm called ReducedBasisFromParameters, which describes the mapping $\Phi_{\boldsymbol{E}} : \mathbb{K}^N \to \mathcal{C}(\boldsymbol{E})$. This procedure is rather straightforward;

the algorithm for the inverse operation, called ParametersFromReducedBasis, is slightly more involved, and is described in the next section. We still use the notation of Section 4.3.2, writing in particular $M_i \in \mathbb{K}[x]$ for the polynomial coefficient of y^{n_i} in both g_i and h_i, for all i, and m_i for its degree.

We compute the h_i's, and then the g_i's, in descending order. To obtain the former, we simply use Eq. (4.3.0.2). For any $i = s-1, \ldots, 0$, assuming we know h_i and $g_{i+1}, \ldots, g_s$, let us show how to obtain g_i by reducing h_i (for $i = s$, we have $g_s = h_s$), using procedure Reduction from the previous section.

Using Euclidean division with respect to x, the polynomial h_i can be written as $h_i = A_i M_{i+1} + B_i$, with A_i and B_i in $\mathbb{K}[x, y]$ and $\deg(B_i, x) < m_{i+1}$.

Recall now that all polynomials $g_{i+1}, \ldots, g_s$ are multiples of M_{i+1}, and that the family $\mathcal{G}_i = (g_{i+1}/M_{i+1}, \ldots, g_s/M_{i+1})$ is a zero-dimensional Gröbner basis (as pointed out after Eq. (4.3.0.1)). Set $\bar{h}_i = (A_i \text{ rem } \mathcal{G}_i)M_{i+1} + B_i$; we claim that $\bar{h}_i = g_i$. First, we determine its initial term: all monomials in A_i rem $\mathcal{G}_i$, and thus in $(A_i \text{ rem } \mathcal{G}_i)M_{i+1}$, have y-degree less than n_{i+1}, whereas B_i contains the initial term $x^{m_i}y^{n_i}$ of h_i. Thus the initial term of $\bar{h}_i$ is still $x^{m_i}y^{n_i}$. Next, we verify that $\bar{h}_i$ is reduced modulo $g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_s$.

- None of $g_1, \ldots, g_{i-1}$ can reduce any term in $\bar{h}_i$, since this polynomial has y-degree n_i.
- Since A_i rem $\mathcal{G}_i$ is reduced modulo $\mathcal{G}_i$, $(A_i \text{ rem } \mathcal{G}_i)M_{i+1}$ is reduced modulo $g_{i+1}, \ldots, g_s$.
- Since B_i has x-degree less than m_{i+1}, it is also reduced modulo $g_{i+1}, \ldots, g_s$.

The last observation is that the difference $\bar{h}_i - h_i$ is in the ideal $\langle g_{i+1}, \ldots, g_s \rangle$. Altogether, this establishes $\bar{h}_i = g_i$.

With this, we can give our algorithm to compute $g_0, \ldots, g_s$. For the reduction of the bivariate polynomial A_i modulo $\mathcal{G}_i$, we use our procedure Reduction. Note that the degree assumptions for that procedure are satisfied: the polynomial A_i has x-degree less than $m_s - m_{i+1}$ and y-degree less than n_{i+1}, which are precisely the maximal x-degrees and y-degrees of the elements in $\mathcal{G}_i$.

As before, we assume that given $\boldsymbol{E}$, we can directly access the integers s, $(d_i)_{1 \le i \le s}$ and $(e_i)_{1 \le i \le s}$ and use them freely in the pseudo-code.

Algorithm 4.5.3 ReducedBasisFromParameters($\boldsymbol{E}, (\lambda_1, \ldots, \lambda_N)$)

```
Input: monomials $\boldsymbol{E}$, $(\lambda_1, \ldots, \lambda_N)$ in $\mathbb{K}^N$
Output: the reduced Gröbner basis of $\Phi_{\boldsymbol{E}}(\lambda_1, \ldots, \lambda_N)$
$(\sigma_{i,j})_{i,j} \leftarrow$ SigmaFromParameters($\boldsymbol{E}, (\lambda_1, \ldots, \lambda_N)$)
$M_0 \leftarrow 1$
for $i = 1, \ldots, s$ do $M_i \leftarrow (x^{d_i} - \sigma_{i-1,i-1})M_{i-1}$
$h_s \leftarrow M_s$; $g_s \leftarrow M_s$
for $i = 0, \ldots, s-1$ do
    $T_i \leftarrow$ KroneckerMultiply($y^{e_{i+1}}, h_{i+1}$) $+ \cdots +$ KroneckerMultiply($\sigma_{s,i}, h_s$)
    $h_i \leftarrow T_i$ div $(x^{d_{i+1}} - \sigma_{i,i})$
    $\mathcal{G}_i \leftarrow (g_{i+1}$ div $M_{i+1}, \ldots, g_s$ div $M_{i+1})$
    $A_i, B_i \leftarrow h_i$ div $M_{i+1}, h_i$ rem $M_{i+1}$
    $\bar{A}_i \leftarrow$ Reduction($A_i, \mathcal{G}_i$)
    $g_i \leftarrow \bar{A}_i M_{i+1} + B_i$
return $(g_0, \ldots, g_s)$
```

Proposition 4.5.1. *Given monomials* $\boldsymbol{E}$ *and* $(\lambda_1, \ldots, \lambda_N)$ *in* $\mathbb{K}$, *ReducedBasisFromParameters*$(\boldsymbol{E}, (\lambda_1, \ldots, \lambda_N))$ *returns the reduced Gröbner basis of* $\Phi_{\boldsymbol{E}}(\lambda_1, \ldots, \lambda_N)$ *using* $\tilde{O}(s^2 n_0 m_s)$ *operations* $(+, \times)$ *in* $\mathbb{K}$.

Proof. Correctness follows from the previous discussion. Regarding the runtime, the first step does no arithmetic operation, and computing each polynomial M_i takes $\tilde{O}(\delta)$ operations, for a total of $\tilde{O}(s\delta)$.

For a given index i, computing T_i involves at most s polynomial multiplications, each of which uses $\tilde{O}(n_0 m_s)$ operations $(+, \times)$ in $\mathbb{K}$; we can deduce h_i in the same asymptotic time. The Euclidean divisions needed to compute $\mathcal{G}_i$ cost $\tilde{O}(s\delta)$ operations (since all polynomials in $\mathcal{G}$ are supported on an initial segment of size δ), and the one for A_i and B_i costs $\tilde{O}(n_0 m_s)$, for the same reason. Proposition 4.4.3 shows that we compute $\bar{A}_i$ in $\tilde{O}(n_0 m_s + s\delta)$ operations $(+, \times)$. Finally, the product and sum giving g_i take $\tilde{O}(n_0 m_s)$ operations $(+, \times)$ as well.

Altogether, the cost at step i is $\tilde{O}(s n_0 m_s + s\delta)$, which is $\tilde{O}(s n_0 m_s)$, and the overall runtime estimate follows. □

It will be useful to note that the algorithm does not perform divisions, so if the input parameters lie in a ring $\mathbb{A} \subset \mathbb{K}$, the output polynomials $\mathcal{G}$ all have coefficients in $\mathbb{A}$.

The whole procedure can be adapted to deal with punctual Gröbner cells in a straightforward manner, by using SigmaFromPunctualParameters at Line 1. The resulting function is called ReducedBasisFromPunctualParameters, and features a similar runtime.

4.6 Computing the Gröbner parameters

We can now give our algorithms to compute the Gröbner parameters of a zero-dimensional ideal I.

We do this in two different contexts. The first situation is the recovery of these parameters starting from the reduced Gröbner basis of I (*i.e.*, computing the map $\Phi_{\boldsymbol{E}}^{-1}$ defined in the previous sections). This is relatively straightforward, using a sequence of Euclidean divisions.

The second variant we present is the core ingredient of our main algorithm. Here, we consider an ideal J given by generators $f_1, \ldots, f_t$, a zero-dimensional ideal I containing J, and we describe a system of polynomials which admits the Gröbner parameters of I as a solution with multiplicity one. In that, we follow previous work of Hauenstein, Mourrain, Szanto [88] that was in the context of border bases representations.

These latter equations are in general too complex to be dealt with directly. In the next section, we will use them to describe our main algorithm, a version of Newton iteration to compute the Gröbner parameters of I as above.

4.6.1 Starting from a reduced basis

In this subsection, we assume that we are given the reduced Gröbner basis $\mathcal{G} = (g_0, \ldots, g_s)$ of a zero-dimensional ideal I, and we show how to compute its Gröbner parameters. We also indicate how the procedure simplifies slightly when I is $\langle x, y\rangle$-primary.

Our notation is as before: the initial terms of the polynomials $(g_0, \ldots, g_s)$ are written $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$, the degree of $\mathcal{G}$ is δ and $N = \delta + m_s$ is the number of Gröbner parameters. In what follows, we compute the polynomials $(\sigma_{i,j})_{i,j}$ appearing in the syzygies (4.3.0.2), whose coefficients are the Gröbner parameters of I. Recall that we write $D_i = x^{d_i} - \sigma_{i-1,i-1}$ for $i = 1, \ldots, s$, and $M_i = (x^{d_1} - \sigma_{0,0}) \cdots (x^{d_i} - \sigma_{i-1,i-1})$ for $i = 0, \ldots, s$, with the empty product being equal to 1.

4.6.1.1 Deriving the algorithm Knowing the reduced Gröbner basis $\mathcal{G} = (g_0, \ldots, g_s)$, some of the polynomials $(\sigma_{i,j})$ are easy to compute: for $i = 1, \ldots, s$, we saw in the previous section that the polynomial coefficient of y^{n_i} in g_i is none other than M_i. Knowing $M_1, \ldots, M_s$ gives us $D_1, \ldots, D_s$, and thus $\sigma_{0,0}, \ldots, \sigma_{s-1,s-1}$, by successive divisions.

Let now $h_0, \ldots, h_s$ be the non-reduced Gröbner basis already used previously, that satisfies Eq. (4.3.0.2), and recall that for $i = 0, \ldots, s$, M_i divides h_i. We define $H_i = h_i/M_i$, and consider again Eq. (4.3.0.3), which is a rewriting of (4.3.0.2):

$$D_{i+1}h_i - y^{e_{i+1}}h_{i+1} = \sum_{j=i+1}^{s} \sigma_{j,i}h_j,$$

in which both left- and right-hand sides can be divided by M_{i+1}. Carrying out the division, we obtain

$$H_i - y^{e_{i+1}}H_{i+1} = \sum_{j=i+1}^{s} \sigma_{j,i}D_{i+2}\cdots D_jH_j. \tag{4.6.0.1}$$

Fix i in $\{0, \ldots, s-1\}$, and assume that we have computed $H_{i+1}, \ldots, H_s$; we show how to compute $\sigma_{i+1,i}, \ldots, \sigma_{s,i}$, and then H_i.

By construction, the polynomials $(g_0, \ldots, g_i, h_{i+1}, \ldots, h_s)$ also form a minimal Gröbner basis of I. The polynomial $h_i - g_i$ is in I, so it reduces to zero through division by these polynomials. Since g_i and h_i both have M_i as polynomial coefficient of y^{n_i}, $h_i - g_i$ has degree less than n_i in y. This implies that the only polynomials in the list that can reduce it are $h_{i+1}, \ldots, h_s$. We reduce $h_i - g_i$ by h_{i+1}, then h_{i+2}, etc, in this order; for $j = i, \ldots, s$, write $R_{i,j}$ for the remainder obtained after reduction by $h_{i+1}, \ldots, h_j$, so that $R_{i,i} = h_i - g_i$.

Lemma 4.6.1. *For $j = i, \ldots, s$, $R_{i,j}$ has y-degree less than n_j.*

Proof. We pointed out that this is true for $j = i$, so we suppose that the claim holds for some index $j < s$ and prove it for index $j+1$. To obtain $R_{i,j+1}$, we reduce $R_{i,j}$ by h_{j+1}, which has initial term $x^{m_{j+1}}y^{n_{j+1}}$, so that we can write $R_{i,j+1} = A_{j+1} + B_{j+1}$, with $\deg(B_{j+1}, y) < n_{j+1}$, $\deg(A_{j+1}, x) < m_{j+1}$ and all terms in A_{j+1} having y-degree at least n_{j+1}. To conclude, we prove that $A_{j+1} = 0$.

Since we use the lexicographic order $y \succ x$, reduction of a term by h_{j+1} does not increase its y-degree; since $R_{i,j}$ had y-degree less than n_j by assumption, it is also the case for A_{j+1}. In particular, A_{j+1} is reduced modulo $\mathcal{H}$. Since $R_{i,j}$ reduces to zero modulo $\mathcal{H}$, it follows that $A_{j+1} + (B_{j+1} \text{ rem } \mathcal{H}) = 0$. Now, for the same reason as above, $(B_{j+1} \text{ rem } \mathcal{H})$ has y-degree less than n_{j+1}, so that the supports of A_{j+1} and $(B_{j+1} \text{ rem } \mathcal{H})$ do not overlap. This implies that $A_{j+1} = (B_{j+1} \text{ rem } \mathcal{H}) = 0$, as claimed. □

This lemma shows that the reduction of $h_i - g_i$ induces an equality of the form

$$h_i - g_i = \sum_{j=i+1}^{s} q_{j,i} h_j,$$

for some polynomials $q_{j,i}$ in $\mathbb{K}[x, y]$ satisfying $\deg(q_{j,i}, y) < n_{j-1} - n_j = e_j$ for all j. Equivalently, we may rewrite this as

$$h_i = g_i + \sum_{j=i+1}^{s} q_{j,i} M_j H_j,$$

whence, after dividing by M_i,

$$H_i = G_i + \sum_{j=i+1}^{s} q_{j,i} D_{i+1} \cdots D_j H_j. \tag{4.6.0.2}$$

Combining (4.6.0.1) and (4.6.0.2), we get

$$G_i - y^{e_{i+1}} H_{i+1} = \sum_{j=i+1}^{s} Q_{j,i} H_j, \quad \text{with} \quad Q_{j,i} = (\sigma_{j,i} - q_{j,i} D_{i+1}) D_{i+2} \cdots D_j. \tag{4.6.0.3}$$

Notice in particular that for all j, we have $\deg(Q_{j,i}, y) < e_j$ and thus $\deg(Q_{j,i} H_j, y) < n_{j-1}$.

In this paragraph, for F in $\mathbb{K}[x, y]$, we write $\bar{F}$ for its residue class in $\mathbb{B}[y]$, with $\mathbb{B} = \mathbb{K}[x]/\langle D_{i+1} \cdots D_s \rangle$. Take j in $i+1, \ldots, s-1$ and suppose that we know $\bar{Q}_{i+1,i}, \ldots, \bar{Q}_{j-1,i}$. Split the sum in (4.6.0.3) as $A = Q_{j,i} H_j + R$ with

$$A = G_i - y^{e_{i+1}} H_{i+1} - \sum_{k=i+1}^{j-1} Q_{k,i} H_k \quad \text{and} \quad R = \sum_{k=j+1}^{s} Q_{k,i} H_k.$$

Over $\mathbb{B}[y]$, $\bar{R}$ has degree (in y) less than n_j; since $\bar{H}_j$ is monic of degree n_j, the relation $\bar{A} = \bar{Q}_{j,i} \bar{H}_j + \bar{R}$ describes the Euclidean division of $\bar{A}$, which is known, by $\bar{H}_j$, which is known as well. If we let $Q^*_{i,j}$ be the canonical lift of $\bar{Q}_{i,j}$ to $\mathbb{K}[x, y]$, we obtain

$$\begin{aligned} Q^*_{j,i} &= Q_{j,i} \text{ rem } D_{i+1} \cdots D_s \\ &= (\sigma_{j,i} - q_{j,i} D_{i+1}) D_{i+2} \cdots D_j \text{ rem } D_{i+1} \cdots D_s. \end{aligned}$$

It follows that $Q^*_{i,j}$ is divisible by $D_{i+2} \cdots D_j$, and that

$$Q^*_{i,j} \text{ div } (D_{i+2} \cdots D_j) = (\sigma_{j,i} - q_{j,i} D_{i+1}) \text{ rem } D_{i+1} D_{j+1} \cdots D_s.$$

Since $\deg(\sigma_{j,i}, x) < d_{i+1}$, reducing this modulo D_{i+1} finally gives us $\sigma_{j,i}$. Noticing also that the remainder $\bar{R}$ gives us the next value of $\bar{A}$, we obtain Algorithm ParametersFromReducedBasis.

In the following proposition, in preparation for the discussion in the next subsection, we point out in particular that the algorithm does not perform any division.

Proposition 4.6.1. *Given a minimal reduced Gröbner basis $\mathcal{G} = (g_0, \dots, g_s)$ in $\mathbb{K}[x, y]$, ParametersFromReducedBasis$(\mathcal{G})$ returns the Gröbner coefficients of $\mathcal{G}$ using $O\tilde{~}(s^2 n_0 m_s)$ operations $(+, \times)$ in $\mathbb{K}$.*

Algorithm 4.6.1 ParametersFromReducedBasis($\mathcal{G}$)

```
Input: $\mathcal{G} = (g_0, \dots, g_s)$ in $\mathbb{K}[x, y]^s$
Assumptions: $\mathcal{G}$ is a minimal reduced Gröbner basis, with initial terms
    $(y^{n_0}, \dots, x^{m_s})$ listed in decreasing order
Output: $(\lambda_1, \dots, \lambda_N)$ in $\mathbb{K}^N$
for $i = 0, \dots, s$ do $x^{m_i} y^{n_i} \leftarrow$ InitialTerm$(g_i)$
$M_0 \leftarrow 1$, $G_0 \leftarrow g_0$
for $i = 1, \dots, s$ do
    $M_i \leftarrow$ PolynomialCoefficient$(g_i, y^{n_i})$          ▷ $M_i$ monic in $\mathbb{K}[x]$
    $G_i \leftarrow g_i$ div $M_i$
    $D_i \leftarrow M_i$ div $M_{i-1}$          ▷ $D_i$ monic in $\mathbb{K}[x]$
    $n_{i-1,i-1} \leftarrow x^{d_i} - D_i$          ▷ $d_i = m_i - m_{i-1}$
$H_s \leftarrow 1$
for $i = s-1, \dots, 0$ do
    $H_i \leftarrow y^{e_{i+1}} H_{i+1}$          ▷ $e_{i+1} = n_i - n_{i+1}$
    $\bar{A} \leftarrow \bar{G}_i - y^{e_{i+1}} \bar{H}_{i+1}$     ▷ computation done in $\mathbb{B}[y]$, with $\mathbb{B} = \mathbb{K}[x]/\langle D_{i+1} \cdots D_s \rangle$
    for $j = i+1, \dots, s$ do
        $\bar{Q}_{j,i} \leftarrow \bar{A}$ div $\bar{H}_j$, $\bar{A} \leftarrow \bar{A}$ rem $\bar{H}_j$          ▷ Euclidean division done in $\mathbb{B}[y]$
        $Q^*_{j,i} \leftarrow$ canonical lift of $\bar{Q}_{j,i}$ to $\mathbb{K}[x, y]$
        $\sigma_{j,i} \leftarrow (Q^*_{j,i}$ div $D_{i+2} \cdots D_j)$ rem $D_{i+1}$
        $H_i \leftarrow H_i +$ KroneckerMultiply$(\sigma_{j,i}, D_{i+2} \cdots D_j H_j)$
return ParametersFromSigma$((y^{n_0}, \dots, x^{m_s}), (\sigma_{j,i})_{0 \le i \le s-1, i \le j \le s})$
```

As before, the modifications needed to deal with the punctual Gröbner cell are elementary; it suffices to invoke PunctualParametersFromSigma at the last step. The resulting procedure will be written PunctualParametersFromReducedBasis. Before proving the proposition, we give an example of computation of punctual Gröbner coefficients.

❧ Example 4.6.1

Given $\mathcal{G}$ as in the introduction from Example 4.1.1,

$$\begin{aligned}&y^4 + \tfrac{17}{14}xy - \tfrac{17}{7}x^2,\\ &xy^3 - \tfrac{10}{9}x^3,\\ &x^2y - 2x^3,\\ &x^4,\end{aligned}$$

Algorithm PunctualParametersFromReducedBasis computes

$$\begin{aligned}\sigma_{0,0} = \sigma_{1,0} = 0, \quad \sigma_{2,0} = \frac{17}{14}, \quad \sigma_{3,0} = \frac{40}{9},\\ \sigma_{1,1} = \sigma_{2,1} = 0, \quad \sigma_{3,1} = -\frac{10}{9},\\ \sigma_{2,2} = 0, \quad \sigma_{3,2} = -2x\end{aligned}$$

and thus

$$\lambda_1 = 0, \quad \lambda_2 = \frac{17}{14}, \quad \lambda_3 = \frac{40}{9}, \quad \lambda_4 = -\frac{10}{9}, \quad \lambda_5 = -2. \tag{4.6.0.4}$$

Proof. We already established correctness of the algorithm. By inspection, we see that all steps involve only additions and multiplications in $\mathbb{K}$, using only integer constants, since all that is done are multiplications or Euclidean divisions by monic polynomials, either in $\mathbb{K}[x]$ or in $\mathbb{B}[y]$, with $\mathbb{B}$ of the form $\mathbb{K}[x]/\langle D_{i+1} \cdots D_s\rangle$ (this in turn reduces to additions and multiplications in $\mathbb{K}$).

It remains to establish the runtime of the algorithm. Each pass in the loop at Line 3 uses $O\tilde{~}(\delta)$ operations $(+, \times)$, for a total of $O\tilde{~}(s\delta)$. To continue the analysis, we first note that for all i, the polynomial H_i computed by the algorithm has x-degree less than $d_{i+1} + \cdots + d_s$, which is less than m_s, and y-degree n_i. The same bounds holds for $\deg(Q^*_{j,i}, x)$ (by construction); the y-degree of this polynomial is less than e_j, as mentioned during the derivation of the algorithm.

Since G_i satisfies the same degree bound $\deg(G_i, x) < d_{i+1} + \cdots + d_s$ as H_i, the reduction of $G_i - y^{e_{i+1}}H_{i+1}$ modulo $D_{i+1} \cdots D_s$ at is free. At each pass through

Line 13, the Euclidean division takes $O\tilde{~}(n_{j-1}) \subset O\tilde{~}(n_0)$ operations $(+, \times)$ in $\mathbb{B}$, which is $O\tilde{~}(n_0 m_s)$ operations $(+, \times)$ in $\mathbb{K}$. The degree bounds given above show that the cost of computing $\sigma_{j,i}$ and updating H_i admits the same upper bound $O\tilde{~}(n_0 m_s)$. Since we enter the inner For loop at Line 12 $O(s^2)$ times, this gives a total cost $O\tilde{~}(s^2 n_0 m_s)$. □

Let us now see how to formalize the observation that the coefficients computed by Algorithm ParametersFromReducedBasis are polynomial expressions of the coefficients of $\mathcal{G}$.

Assume that the terms $\boldsymbol{E}$ are fixed, let $\mu_1, \ldots, \mu_\delta$ be the monomials not in $\langle \boldsymbol{E} \rangle$, ordered in an arbitrary fashion, and let $\Gamma_{0,1}, \ldots, \Gamma_{s,\delta}$ be $(s+1)\delta$ new variables over $\mathbb{Z}$. We set $\mathbb{A}_{\boldsymbol{E}} = \mathbb{Z}[\Gamma_{0,1}, \ldots, \Gamma_{s,\delta}]$.

Because the algorithm only performs additions and multiplications, and uses constants from the image of $\mathbb{Z}$ in $\mathbb{K}$, we deduce that there exist $P_{1,\boldsymbol{E}}, \ldots, P_{N,\boldsymbol{E}}$ in $\mathbb{A}_{\boldsymbol{E}} = \mathbb{Z}[\Gamma_{0,1}, \ldots, \Gamma_{s,\delta}]$ such that given any reduced Gröbner basis $\mathcal{G} = (g_0, \ldots, g_s)$ with initial terms $\boldsymbol{E}$ and with coefficients in $\mathbb{K}$ (or any extension of it, as we choose below), the Gröbner parameters of $\mathcal{G}$ are obtained by evaluating $P_{1,\boldsymbol{E}}, \ldots, P_{N,\boldsymbol{E}}$ at the coefficients of $\mathcal{G}$.

Correctness of the algorithm can then be restated as follows. Let $\Lambda_1, \ldots, \Lambda_N$ be another set of new variables over $\mathbb{K}$, that stand for "generic" Gröbner parameters, and define $\mathbb{L} = \mathbb{K}(\Lambda_1, \ldots, \Lambda_N)$. Let further $g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}}$ be the polynomials obtained as output of ReducedBasisFromParameters$(\boldsymbol{E}, (\Lambda_1, \ldots, \Lambda_N))$. Since that algorithm as well performs only additions and subtractions (Proposition 4.5.1), these polynomials actually have coefficients in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N] \subset \mathbb{L}$. For $i = 0, \ldots, s$ and $j = 1, \ldots, \delta$, let then $R_{i,j} \in \mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$ be the coefficient of the monomial μ_j in $g_{i,\mathbb{L}}$. We deduce from our discussion that $P_{i,\boldsymbol{E}}(R_{0,1}, \ldots, R_{s,\delta}) = \Lambda_i$ holds for all i. We will use this observation in the next subsection.

4.6.2 Polynomial equations for the Gröbner parameters

Let now $f_1, \ldots, f_t$ be polynomials in $\mathbb{K}[x, y]$; in this subsection, those are our inputs, and we denote by J the ideal they generate in $\mathbb{K}[x, y]$. Let further I be an ideal in $\mathbb{K}[x, y]$ such that the following properties hold:

A_1. I has dimension zero;

A_2. there exists an ideal $I' \subset \mathbb{K}[x, y]$ such that $I + I' = \langle 1 \rangle$ and $II' = J$.

Equivalently, I is the intersection (or product) of some zero-dimensional primary components of J. This is for instance the case if the origin $(0, 0)$ is isolated in $V(J)$ and I is the $\langle x, y \rangle$-primary component of J, or if $I = J$ and $V(J)$ is finite.

Let $\mathcal{G} = (g_0, \ldots, g_s) \subset \mathbb{K}[x, y]$ be the reduced lexicographic Gröbner basis of I. We denote by $\boldsymbol{E}$ the initial terms of the polynomials in $\mathcal{G}$, written as before as

$$\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s}).$$

In what follows, we assume that $\boldsymbol{E}$ is known, but not $\mathcal{G}$; we show how to recover the Gröbner parameters of I (and thus $\mathcal{G}$ itself).

We let δ be the degree of I, and $\mu_1, \ldots, \mu_\delta$ be the monomials not in $\langle \boldsymbol{E} \rangle$, ordered in an arbitrary way. Let further $N = \delta + m_s$ be the number of parameters for the Gröbner cell $\mathcal{C}(\boldsymbol{E})$, and let $(\lambda_1, \ldots, \lambda_N) = \phi_{\boldsymbol{E}}^{-1}(I) \in \mathbb{K}^N$ be the Gröbner parameters associated to I. In this subsection, we define a system of $t\delta$ equations $\mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta}$ in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$, where $\Lambda_1, \ldots, \Lambda_N$ are new variables, and we prove that $(\lambda_1, \ldots, \lambda_N)$ is a solution of multiplicity 1 to these equations.

As in the previous subsection, let $\mathbb{L} = \mathbb{K}(\Lambda_1, \ldots, \Lambda_N)$ and let $g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}}$ be the parametric Gröbner basis of $\mathcal{C}(\boldsymbol{E})$ given by ReducedBasisFromParameters$(\boldsymbol{E}, (\Lambda_1, \ldots, \Lambda_N))$. Recall that all polynomials $g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}}$ have coefficients in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$; this implies in particular that for A in $\mathbb{K}[x, y]$, the remainder A rem $\langle g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}} \rangle$ is in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N][x, y]$. For $j = 1, \ldots, \delta$, we then denote by $\mathcal{N}_i$ the following $\mathbb{K}$-linear map:

$$\begin{aligned} \mathcal{N}_j : \mathbb{K}[x, y] &\to \mathbb{K}[\Lambda_1, \ldots, \Lambda_N] \\ A &\mapsto \mathrm{coeff}(A \text{ rem } \langle g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}} \rangle, \mu_j), \end{aligned}$$

with $\mu_1, \ldots, \mu_\delta$ the monomials not in $\langle \boldsymbol{E} \rangle$, as defined above. For $i = 1, \ldots, t$, we then let

$$\mathscr{E}_{i,1}, \ldots, \mathscr{E}_{i,\delta} = \mathcal{N}_1(f_i), \ldots, \mathcal{N}_\delta(f_i),$$

thus defining $t\delta$ polynomials $\mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta}$ in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$. The following key property for these equations was inspired by [88, Theorem 4.8], which was stated in the context of border bases.

Proposition 4.6.2. $(\lambda_1, \ldots, \lambda_N)$ *is a solution of* $\mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta}$ *of multiplicity* 1.

Proof. Let $\mathcal{I}$ be the ideal generated by all polynomials $\mathcal{N}_i(g_j)$, for $i = 1, \ldots, \delta$ and $j = 0, \ldots, s$, and let $R_{0,1}, \ldots, R_{s,\delta} \in \mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$ be the coefficients of $(g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}})$, as in the previous subsection. Then, for $i = 1, \ldots, \delta$ and $j = 0, \ldots, s$, the polynomial $\mathcal{N}_i(g_j)$ is equal to $R_{j,i}(\lambda_1, \ldots, \lambda_N) - R_{j,i}$. In particular, $(\lambda_1, \ldots, \lambda_N)$ is in the zero-set of $\mathcal{I}$.

Recall further from the previous subsection the existence of polynomials $P_{1,\boldsymbol{E}}, \ldots, P_{N,\boldsymbol{E}}$, with $P_{k,\boldsymbol{E}}(R_{0,1}, \ldots, R_{s,\delta}) = \Lambda_k$ for all k. The fact that $R_{j,i}(\lambda_1, \ldots, \lambda_N) - R_{j,i}$ is in $\mathcal{I}$ for all i, j implies that

$$P_{k,\boldsymbol{E}}(R_{0,1}(\lambda_1, \ldots, \lambda_N), \ldots, R_{s,\delta}(\lambda_1, \ldots, \lambda_N)) - P_{k,\boldsymbol{E}}(R_{0,1}, \ldots, R_{s,\delta})$$

is in $\mathcal{I}$ as well, for all $k = 1, \ldots, N$. The left-hand side is λ_k, and the right-hand side Λ_k, so that $\mathcal{I}$ contains all polynomials $\Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N$. Taken together, the two paragraphs so far establish that $\mathcal{I} = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle$.

Let now $\mathscr{J}$ be the ideal generated in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$ by the polynomials $\mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta}$. Remark first that for any $a, b \geq 0$ and $i = 1, \ldots, t$,

$$(x^a y^b f_i) \text{ rem } \langle g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}} \rangle = \sum_{j=1}^{\delta} \mathscr{N}_j(f_i)(x^a y^b \mu_j \text{ rem } \langle g_{0,\mathbb{L}}, \ldots, g_{s,\mathbb{L}} \rangle).$$

It follows that for any A in $J = \langle f_1, \ldots, f_t \rangle$, and for $j = 1, \ldots, \delta$, $\mathscr{N}_j(A)$ is in $\mathscr{J}$. For the same reason, for A in $I = \langle g_0, \ldots, g_s \rangle$, and for $j = 1, \ldots, \delta$, $\mathscr{N}_j(A)$ is in $\mathcal{I}$. We will also need the fact that for A in I^2, and for all j, $\mathscr{N}_j(A)$ is in $\mathcal{I}^2$; this is established similarly.

Recall now our second assumption on I': there exists an ideal $I' \subset \mathbb{K}[x, y]$ such that $I + I' = \langle 1 \rangle$ and $II' = J$. Since J is contained in I, the statements in the previous paragraph imply that $\mathscr{J}$ is contained in $\mathcal{I} = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle$, so that $(\lambda_1, \ldots, \lambda_N)$ is in the zero-set of $\mathscr{J}$. This proves the first claim of the proposition.

Let further K, K' be in resp. I and I' such that $K + K' = 1$. For $i = 0, \ldots, s$, g_i is in I, so that $g_i K' = g_i - g_i K$ is in $II' = J$. By the remark above, for $j = 1, \ldots, \delta$, $A_{j,i} := \mathscr{N}_j(g_i) - \mathscr{N}_j(g_i K)$ is then in $\mathscr{J}$, whereas $\mathscr{N}_j(g_i K)$ is in $\mathcal{I}^2$.

Consider the Jacobian matrix $\boldsymbol{J}$ of all polynomials $A_{j,i}$ at $(\lambda_1, \ldots, \lambda_N)$. Because all terms $\mathscr{N}_j(g_i K)$ are in $\mathcal{I}^2 = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle^2$, their Jacobian matrix vanishes at $(\lambda_1, \ldots, \lambda_N)$, so that $\boldsymbol{J}$ is simply the Jacobian matrix of the polynomials $\mathscr{N}_j(g_i)$ at $(\lambda_1, \ldots, \lambda_N)$. Because these polynomials generate the ideal $\mathcal{I} = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle$, this matrix has trivial kernel. Thus, $\mathscr{J}$ has multiplicity 1 at $(\lambda_1, \ldots, \lambda_N)$. □

In the particular case where $I = J$, we have a slightly stronger result.

Corollary 4.6.1. *Suppose that $I = \langle f_1, \ldots, f_t \rangle$. Then, $\langle \mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta} \rangle = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle$ in $\mathbb{K}[\Lambda_1, \ldots, \Lambda_N]$.*

Proof. Using the notation in the proof of the proposition, we see that if $I = J$, then $\mathcal{I} = \mathscr{J}$, and we proved that $\mathcal{I} = \langle \Lambda_1 - \lambda_1, \ldots, \Lambda_N - \lambda_N \rangle$. □

In our other particular case, where I is the $\langle x, y \rangle$-primary component of J, we can obtain a similar stronger statement. Recall that the punctual Gröbner cell $\mathcal{C}_0(\boldsymbol{E})$ has dimension $N' = \delta - n_0$, and that the parameters for $\mathcal{C}_0(\boldsymbol{E})$ are obtained by setting $N - N'$ parameters to zero in the parameters $\Lambda_1, \ldots, \Lambda_N$ of $\mathcal{C}(\boldsymbol{E})$.

Let $\tau_1, \ldots, \tau_{N-N'}$ be the indices of these parameters set to zero, and let $\Lambda_{\sigma_1}, \ldots, \Lambda_{\sigma_{N'}}$ be the remaining N' parameters. For $i = 1, \ldots, t$ and $j = 1, \ldots, \delta$, let $\mathscr{F}_{i,j}$ be the polynomial in $\mathbb{K}[\Lambda_{\sigma_1}, \dots, \Lambda_{\sigma_{N'}}]$ obtained by setting $\Lambda_{\tau_1}, \dots, \Lambda_{\tau_{N-N'}}$ to zero in $\mathscr{E}_{i,j}$. Then, we have the following.

Corollary 4.6.2. *Suppose that I is $\langle x, y\rangle$-primary. Then, $\langle \mathscr{F}_{1,1}, \dots, \mathscr{F}_{t,\delta}\rangle = \langle \Lambda_{\sigma_1} - \lambda_{\sigma_1}, \dots, \Lambda_{\sigma_{N'}} - \lambda_{\sigma_{N'}}\rangle$ in $\mathbb{K}[\Lambda_{\sigma_1}, \dots, \Lambda_{\sigma_{N'}}]$.*

Proof. We proved in Proposition 4.6.2 that $\lambda_{\sigma_1}, \dots, \lambda_{\sigma_{N'}}$ is a solution of $\mathscr{F}_{1,1}, \dots, \mathscr{F}_{t,\delta}$. Besides, since the Jacobian matrix of $\mathscr{E}_{1,1}, \dots, \mathscr{E}_{t,\delta}$ has trivial kernel at $(\lambda_1, \dots, \lambda_N)$ (with thus $\lambda_{\tau_1} = \cdots = \lambda_{\tau_{N-N'}} = 0$), it is also the case for that of $\mathscr{F}_{1,1}, \dots, \mathscr{F}_{t,\delta}$ at $(\lambda_{\sigma_1}, \dots, \lambda_{\sigma_{N'}})$. The only missing property is thus that $(\lambda_{\sigma_1}, \dots, \lambda_{\sigma_{N'}})$ is the only common solution to these equations. Let $(\lambda^\star_{\sigma_1}, \dots, \lambda^\star_{\sigma_{N'}}) \in \overline{\mathbb{K}}^{N'}$ be such a solution, let $\mathcal{G}^\star$ be the corresponding reduced Gröbner basis, and let $I^\star$ be the ideal it generates (in particular, $V(I^\star) = \{(0,0)\}$). Since by assumption $\mathcal{G}^\star$ reduces $f_1, \dots, f_t$ to zero, we have $J \subset I^\star$.

By assumption on I, there exists an ideal $I' \subset \mathbb{K}[x, y]$ such that $I + I' = \langle 1\rangle$ and $II' = J$. Let K, K' be in resp. I and I' such that $K + K' = 1$; in particular, K' does not vanish at $(0,0)$. Since $V(I^\star) = \{(0,0)\}$, it follows that K' is a unit modulo $I^\star$.

Recall that we write $\mathcal{G} = (g_0, \dots, g_s)$ for the reduced lexicographic Gröbner basis of I. Then, for $i = 0, \dots, s$, the polynomial g_iK' is in II', so in J, and thus in $I^\star$. Since K' is a unit modulo $I^\star$, this means that g_i is in $I^\star$. Altogether, this proves that I is contained in $I^\star$. Since these ideals have the same initial ideals for the lexicographic order, they are then equal. This in turn proves that $(\lambda_{\sigma_1}, \dots, \lambda_{\sigma_{N'}}) = (\lambda^\star_{\sigma_1}, \dots, \lambda^\star_{\sigma_{N'}})$. □

4.6.2.1 Example. In our running example, we consider only the punctual Gröbner cell, and we take f_1 and f_2 as in the introduction. To write the equations for the punctual Gröbner parameters, we consider $g_{0,\mathbb{L}}, \dots, g_{3,\mathbb{L}}$ and set to zero the parameters written $\Lambda_{\tau_1}, \dots, \Lambda_{\tau_{N-N'}}$ above; the resulting polynomials were given in (4.3.0.4), written in variables $\lambda_1, \dots, \lambda_5$ (recall that $N' = 5$ here). After reducing f_1 and f_2 by these polynomials, and taking coefficients (we discard those that are identically zero), we obtain

$$14\Lambda_1, \quad 14\Lambda_2 - 17, \quad -14\Lambda_1\Lambda_5^2 + 14\Lambda_3 - 28\Lambda_4\Lambda_5, \quad 14\Lambda_2\Lambda_5 + 34, \quad -18\Lambda_4 + 10\Lambda_5. \tag{4.6.0.5}$$

As claimed, these polynomials generate the maximal ideal

$$\Lambda_1, \quad \Lambda_2 - 17/14, \quad \Lambda_3 - 40/9, \quad \Lambda_4 + 10/9, \quad \Lambda_5 + 2.$$

Because the input f_1, f_2 and $\mathcal{G}$ have rather small degrees, the equations in (4.6.0.5) can be solved by hand. This is of course not the case in general (although on many examples, several of the equations are indeed linear).

4.7 Newton iteration

We can finally describe our main algorithm, which computes Gröbner parameters using Newton iteration. For this, we will suppose that $\mathbb{K}$ is the field of fractions of a domain $\mathbb{A}$, and we consider a maximal ideal $\mathfrak{m}$ in $\mathbb{A}$, with residual field $\mathbb{k} = \mathbb{A}/\mathfrak{m}$.

Consider the following objects: polynomials $(f_1, \dots, f_t)$ in $\mathbb{A}[x, y]$ and a reduced Gröbner basis $\mathcal{G}$ in $\mathbb{K}[x, y]$. We make the following assumptions:

A_1'. the ideal generated by $\mathcal{G}$ in $\mathbb{K}[x, y]$ is the intersection of some of the primary components of $\langle f_1, \dots, f_t \rangle$,

A_2'. all polynomials in $\mathcal{G}$ are in $\mathbb{A}_\mathfrak{m}[x, y]$,

A_3'. the ideal generated by $\mathcal{G}_\mathfrak{m} = \mathcal{G}$ rem $\mathfrak{m}$ in $\mathbb{k}[x, y]$ is the intersection of some of the primary components of the ideal $\langle f_1 \text{ rem } \mathfrak{m}, \dots, f_t \text{ rem } \mathfrak{m} \rangle$.

The last two items express that $\mathfrak{m}$ is good for $\mathcal{G}$, in the sense of Definition 26. Important cases where the first and third assumptions are satisfied are as in the previous subsection, viz. when $\mathcal{G}_\mathfrak{m}$ and $\mathcal{G}$ generate the ideals $\langle f_1 \text{ rem } \mathfrak{m}, \dots, f_t \text{ rem } \mathfrak{m} \rangle$, resp. $\langle f_1, \dots, f_t \rangle$ themselves, or when they describe the $\langle x, y \rangle$-primary components of these ideals.

Given $\mathfrak{m}$, $(f_1, \dots, f_t)$ and $\mathcal{G}_\mathfrak{m}$, we show here how to compute $\mathcal{G}$ rem $\mathfrak{m}^K$, for an arbitrary $K \geq 1$.

Algorithm LiftOneStep describes the core lifting procedure; it takes as input the Gröbner parameters of $\mathcal{G}$, known modulo $\mathfrak{m}^\kappa$, for some $\kappa \geq 0$, and returns these parameters modulo $\mathfrak{m}^{2\kappa}$ (note that since $\mathcal{G}$ has coefficients in $\mathbb{A}_\mathfrak{m}$ by A_2', its Gröbner parameters are in $\mathbb{A}_\mathfrak{m}$ as well, so reducing them modulo powers of $\mathfrak{m}$ makes sense).

The algorithm simply applies Newton's iteration to the equations $\mathscr{E}_{i,j}$ introduced in the previous subsection. Note however that we never explicitly write down these equations, as they may involve a large number of terms: instead, we only compute their first order Taylor expansions, as this is enough to conduct the iteration. This explains why below, we work modulo the ideal $\langle \Lambda_1, \dots, \Lambda_N \rangle^2$.

Since we want to give a cost estimate that counts operations in $\mathbb{A}_{2\kappa}$, we here assume that we already know the reductions of the input equations $f_1, \dots, f_t$ modulo $\mathfrak{m}^{2\kappa}$; they are written $f_1', \dots, f_t' \in \mathbb{A}_{2\kappa}[x, y]$. Some steps in the algorithm require a few further comments, namely the calls to ReducedBasisFromParameters at Line 5, Reduction at Line 8 and LinearSolve at Line 11.

- At Line 5, we are working with Gröbner parameters written $(\ell_1, \dots, \ell_N)$, that are in $\mathbb{B} = \mathbb{A}_{2\kappa}[\Lambda_1, \dots, \Lambda_N]/\langle \Lambda_1, \dots, \Lambda_N \rangle^2$ (in the algorithm, elements of $\mathbb{B}$

are written as $b_0 + \sum_{i=1}^n b_i \Lambda_i$, for some b_i's in $\mathbb{A}_{2\kappa}$). Recall that Algorithm ReducedBasisFromParameters only does additions and multiplications, and uses constants from $\mathbb{Z}$, so we can run this algorithm with inputs in $\mathbb{B}$; we keep in mind, though, that its properties were only established for inputs in a field.

The same remark applies at Line 8, for Algorithm ReductionGeneralInput.

- The last subroutine solves a linear system over $\mathbb{A}_{2\kappa}$: the inputs are elements of $\mathbb{B}$, which we recall take the form $b_0 + \sum_{i=1}^n b_i \Lambda_i$, for some b_i's in $\mathbb{A}_{2\kappa}$. Procedure LinearSolve then sees these elements are linear equations in the Λ_i's. We will prove that a solution exists, and also that the corresponding matrix admits a maximal minor that does not vanish modulo $\mathfrak{m}$, so that the solution is actually unique.

Algorithm 4.7.1 LiftOneStep($(f'_1, \dots, f'_t), \boldsymbol{E}, (\alpha_1, \dots, \alpha_N)$)

```
Input: $(f'_1, \dots, f'_t)$ in $\mathbb{A}_{2\kappa}[x, y]$, monomials $\boldsymbol{E}$, $(\alpha_1, \dots, \alpha_N)$ in $\mathbb{A}_\kappa^N$
Output: $(\alpha''_1, \dots, \alpha''_N)$ in $\mathbb{A}_{2\kappa}^N$
 1: $(\alpha'_1, \dots, \alpha'_N) \leftarrow$ lift of $(\alpha_1, \dots, \alpha_N)$ to $\mathbb{A}_{2\kappa}^N$
 2: $\mu_1, \dots, \mu_\delta \leftarrow$ monomials not in $\langle \boldsymbol{E} \rangle$
 3: for $i = 1, \dots, N$ do
 4:     $\ell_i \leftarrow \alpha'_i + \Lambda_i$                  ▷ all $\ell_i$ in $\mathbb{B} = \mathbb{A}_{2\kappa}[\Lambda_1, \dots, \Lambda_N]/\langle \Lambda_1, \dots, \Lambda_N \rangle^2$
 5: $\mathcal{G}^* \leftarrow$ ReducedBasisFromParameters($\boldsymbol{E}, (\ell_1, \dots, \ell_N)$)   ▷ computations done over $\mathbb{B}$
 6: $\mathscr{R} \leftarrow [\,]$
 7: for $i = 1, \dots, t$ do
 8:     $r_i \leftarrow$ ReductionGeneralInput($f'_i, \mathcal{G}^*$)      ▷ computations done over $\mathbb{B}$
 9:     for $j = 1, \dots, \delta$ do $r_{i,j} \leftarrow$ coeff($r_i, \mu_j$)      ▷ all $r_{i,j}$ in $\mathbb{B}$
10:     $\mathscr{R} \leftarrow \mathscr{R}$ cat $[r_{i,1}, \dots, r_{i,\delta}]$      ▷ $\mathscr{R}$ is an array with entries in $\mathbb{B}$
11: $(\epsilon_1, \dots, \epsilon_N) \leftarrow$ LinearSolve($\mathscr{R}$)      ▷ all $\epsilon_i$ in $\mathbb{A}_{2\kappa}$
12: for $i = 1, \dots, N$ do $\alpha''_i \leftarrow \alpha'_i + \epsilon_i$      ▷ all $\alpha''_i$ in $\mathbb{A}_{2\kappa}$
13: return $(\alpha''_1, \dots, \alpha''_N)$
```

Proposition 4.7.1. *Suppose that* A'_1, A'_2, A'_3 *hold, and let* $(\lambda_1, \dots, \lambda_N) \in \mathbb{A}_\mathfrak{m}^N$ *be the Gröbner parameters of* $\mathcal{G}$. *Given* $(f_1, \dots, f_t)$ rem $\mathfrak{m}^{2\kappa}$ *and* $(\lambda_1$ rem $\mathfrak{m}^\kappa, \dots, \lambda_N$ rem $\mathfrak{m}^\kappa)$, *Algorithm* *LiftOneStep* *correctly returns* $(\lambda_1$ rem $\mathfrak{m}^{2\kappa}, \dots, \lambda_N$ rem $\mathfrak{m}^{2\kappa})$.

Proof. Let $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{A}_{\mathfrak{m}}^N$ be the Gröbner parameters associated to $\mathcal{G}$. By assumption, the vector $\alpha = (\alpha_1, \ldots, \alpha_N)$ satisfies $\alpha = \lambda \text{ rem } \mathfrak{m}^\kappa$, and the same holds for α'. We prove that the output $\alpha'' = (\alpha_1'', \ldots, \alpha_N'')$ is equal to $\lambda \text{ rem } \mathfrak{m}^{2\kappa}$.

This is simply the classical proof of the validity of Newton's iteration. Let δ be the degree $\mathcal{G}$, and let $\mathscr{E} = (\mathscr{E}_{1,1}, \ldots, \mathscr{E}_{t,\delta})$ be the equations introduced in the previous subsection for the polynomials $f_1, \ldots, f_t$ and $\mathcal{G}$, over the field $\mathbb{K}$. Since all f_i's have coefficients in $\mathbb{A}$, and since the reduction process introduces no new denominator, the polynomials $\mathscr{E}$ are in $\mathbb{A}[\Lambda_1, \ldots, \Lambda_N]$. Using Proposition 4.6.2, assumption A_1' shows that λ is a solution to these equations (and that their Jacobian matrix at λ has trivial kernel, but we will not need this fact directly).

Let further $\mathscr{E}_{\mathfrak{m}} = (\mathscr{E}_{\mathfrak{m},1,1}, \ldots, \mathscr{E}_{\mathfrak{m},t,\delta})$ be these same equations, but this time for the polynomials $f_1 \text{ rem } \mathfrak{m}, \ldots, f_t \text{ rem } \mathfrak{m}$ and $\mathcal{G}_{\mathfrak{m}}$. These are polynomials in $\Bbbk[\Lambda_1, \ldots, \Lambda_N]$, with $\mathscr{E}_{\mathfrak{m}} = \mathscr{E} \text{ rem } \mathfrak{m}$. Using Proposition 4.6.2, assumption A_3' shows that $\lambda \text{ rem } \mathfrak{m}$ is a solution to these equations (which we already could deduce from the previous paragraph) and that their Jacobian matrix at $\lambda \text{ rem } \mathfrak{m}$ has trivial kernel. We will use this below.

We claim that for all i, j, the coefficient $r_{i,j}$ computed at Line 9 is equal to $\mathscr{E}_{i,j}(\ell_1, \ldots, \ell_N)$, computed in $\mathbb{B} = \mathbb{A}_{2\kappa}[\Lambda_1, \ldots, \Lambda_N]/\langle \Lambda_1, \ldots, \Lambda_N \rangle^2$. The only point we have to be careful with is that the output of Algorithm ReducedBasisFromParameters is specified as being a Gröbner basis only if the inputs are in a field. To deal with this, let $\ell_1', \ldots, \ell_N'$ be arbitrary lifts of $\ell_1, \ldots, \ell_N$ to the domain $\mathbb{A}[\Lambda_1, \ldots, \Lambda_N]$, and let $\mathcal{G}'$ be the output of ReducedBasisFromParameters($\boldsymbol{E}, (\ell_1', \ldots, \ell_N')$). These polynomials form a Gröbner basis in $\mathbb{K}(\Lambda_1, \ldots, \Lambda_N)[x, y]$, which happens to have all its coefficients in $\mathbb{A}[\Lambda_1, \ldots, \Lambda_N]$, and $\mathcal{G}^*$ computed at Line 5 is the reduction of $\mathcal{G}'$ modulo $\mathfrak{m}^{2\kappa} + \langle \Lambda_1, \ldots, \Lambda_N \rangle^2$.

Similarly, at Line 8, Algorithm ReductionGeneralInput can take as input polynomials with coefficients in $\mathbb{B}$, but its output was only specified for polynomials with coefficients in a field. This is handled as before, and gives us that for all i, r_i is the reduction modulo $\mathfrak{m}^{2\kappa} + \langle \Lambda_1, \ldots, \Lambda_N \rangle^2$ of the polynomial $f_i \text{ rem } \mathcal{G}'$. Now, the coefficients of $f_i \text{ rem } \mathcal{G}'$ are the polynomials $\mathscr{E}_{i,j}$ evaluated at $(\ell_1', \ldots, \ell_N')$, so altogether, for all i, j, $r_{i,j} = \mathscr{E}_{i,j}(\ell_1, \ldots, \ell_N)$, as an element of $\mathbb{B}$. Taking all i, j at once, we obtain the following equalities over $\mathbb{B}$:

$$\begin{aligned} \mathscr{R} &= \mathscr{E}(\alpha_1' + \Lambda_1, \ldots, \alpha_N' + \Lambda_N) \\ &= \mathscr{E}(\alpha') + \operatorname{jac}(\mathscr{E}, \alpha')[\Lambda_1 \ \cdots \ \Lambda_N]^\top, \end{aligned}$$

where $\operatorname{jac}(\mathscr{E}, \alpha')$ is the Jacobian matrix of $\mathscr{E}$ evaluated at α'. First, we show that the system of linear equations $\mathscr{R}$ has a unique solution $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$ in $\mathbb{A}_{2\kappa}^N$. Indeed, given two solution vectors ϵ and ϵ' in $\mathbb{A}_{2\kappa}^N$, we obtain the relation

$$\mathrm{jac}(\mathscr{E}, \alpha')[\epsilon_1 - \epsilon'_1 \ \cdots \ \epsilon_N - \epsilon'_N]^\top = [0 \ \cdots \ 0]^\top$$

over $\mathbb{A}_{2\kappa}$. We pointed out above that $\mathrm{jac}(\mathscr{E} \text{ rem } \mathfrak{m}, \lambda \text{ rem } \mathfrak{m})$ has trivial kernel, so it admits a non-zero N-minor in $\Bbbk = \mathbb{A}/\mathfrak{m} = \mathbb{A}_{2\kappa}/\mathfrak{m}$. Now, by assumption, $\alpha' \text{ rem } \mathfrak{m} = \lambda \text{ rem } \mathfrak{m}$, so that $\mathrm{jac}(\mathscr{E}, \alpha')$ itself admits an N-minor invertible modulo $\mathfrak{m}$, and thus in $\mathbb{A}_{2\kappa}$. This in turn implies that $\epsilon = \epsilon'$, as vectors over $\mathbb{A}'/\mathfrak{m}^{2\kappa}$. Our first claim is proved.

Second, we show that $\epsilon = (\lambda - \alpha') \text{ rem } \mathfrak{m}^{2\kappa}$ is a solution to these linear equations. Indeed, modulo $\mathfrak{m}^{2\kappa}$, we have the Taylor expansion $\mathscr{E}(\alpha' + \epsilon) = \mathscr{E}(\alpha') + \mathrm{jac}(\mathscr{E}, \alpha')[\epsilon_1 \ \cdots \ \epsilon_N]^\top$: higher-order terms vanish, since all entries of ϵ are by assumption in $\mathfrak{m}^\kappa$. Now, $\alpha' + \epsilon = \lambda \text{ rem } \mathfrak{m}^{2\kappa}$, so $\mathscr{E}(\alpha' + \epsilon) = 0 \text{ rem } \mathfrak{m}^{2\kappa}$, and our claim follows.

The two previous paragraphs prove that at the end of the while loop, the value α'' satisfies $\alpha'' = \alpha' + (\lambda - \alpha') \text{ rem } \mathfrak{m}^{2\kappa} = \lambda \text{ rem } \mathfrak{m}^{2\kappa}$, so the proof is complete. □

Proposition 4.7.2. *Let* $\boldsymbol{E} = (y^{n_0}, x^{m_1}y^{n_1}, \ldots, x^{m_{s-1}}y^{n_{s-1}}, x^{m_s})$ *be the initial terms of* $\mathcal{G}$*, and suppose that all* f_i*'s have degree at most* d*.*

Under assumptions A'_1, A'_2, A'_3, *Algorithm* LiftOneStep *uses* $O\tilde{\ }(s^2 n_0 m_s + t\delta(d^2 + dm_s + s\delta + \delta^{\omega-1}))$ *operations in* $\mathbb{A}_{2\kappa}$*.*

Proof. By convention (see the introduction), lifting each α_i to α'_i takes one operation in $\mathbb{A}_{2\kappa}$, for a total of $O(N) = O(\delta)$ operations.

By Proposition 4.5.1, computing $\mathcal{G}^*$ takes $O\tilde{\ }(s^2 n_0 m_s)$ operations $(+, \times)$ in $\mathbb{B}$, with each such operation taking $O(\delta)$ operations in $\mathbb{A}_{2\kappa}$.

At Line 8, by Proposition 4.4.4, Algorithm ReductionGeneralInput uses $O\tilde{\ }(d^2 + dm_s + n_0 m_s + s\delta)$ operations $(+, \times)$ in $\mathbb{B}$. Here, we know that n_0 is at most d, so the runtime for all f_i's becomes $O\tilde{\ }(t(d^2 + dm_s + s\delta))$ operations in $\mathbb{B}$, which is $O\tilde{\ }(t\delta(d^2 + dm_s + s\delta))$ operations in $\mathbb{A}_{2\kappa}$.

Finally, we have to solve the linear system defined by $\mathscr{R} = 0$ over $\mathbb{A}_{2\kappa}$. This is a system in $t\delta$ equations and N unknowns, and we know that it admits a unique solution in $\mathbb{A}_{2\kappa}^N$, since the corresponding matrix has trivial kernel modulo $\mathfrak{m}$. Even though $\mathbb{A}_{2\kappa}^N$ is not a field, we may still apply fast algorithms, such as the one in [96] (as extended in [99]), replacing zero-tests by invertibility tests; this takes $O\tilde{\ }(t\delta^\omega)$ operations in $\mathbb{A}_{2\kappa}$. □

As usual, if $\mathcal{G}$ (and thus $\mathcal{G}_\mathfrak{m}$) is $\langle x, y\rangle$-primary, we may use a variant of this lifting procedure, called LiftOneStepPunctualParameters, which uses ReducedBasisFromPunctualParameters and PunctualParametersFrom-

ReducedBasis as subroutines. It allows us to work with N' rather than N unknown Gröbner parameters; the proof now relies on Corollary 4.6.2, and the runtime becomes $O\tilde{}(s^2 n_0 m_s + t\delta^2(m_s + \delta^{\omega-2}))$ operations in $\mathbb{A}_{2\kappa}$ (see Proposition 4.4.4).

At this stage, we are almost done with the proof of Theorem 4.1.1: for $K = 2^k$, the algorithm simply computes $\mathcal{G}$ rem $\mathfrak{m}^K$ through repeated calls to Algorithm LiftOneStep. However, this procedure works with Gröbner parameters as input and output. Hence, prior to entering Algorithm LiftOneStep for the first time, we compute the Gröbner parameters of $\mathcal{G}$ rem $\mathfrak{m}$, and after the last call to Algorithm LiftOneStep, we compute $\mathcal{G}$ rem $\mathfrak{m}^K$ using Algorithm ReducedBasisFromParameters. This extra work does not affect the asymptotic runtime, so that we do $O\tilde{}(s^2 n_0 m_s + t\delta(d^2 + dm_s + s\delta))$ operations in $\mathbb{A}/\mathfrak{m}^{2^i}$, for $i = 1, \dots, k$.

The only operations not accounted for so far are the coefficient-wise reductions of the polynomials $f_1, \dots, f_t$ modulo $\mathfrak{m}^2, \dots, \mathfrak{m}^{2^k}$. These cannot be expressed in terms of operations in the residue class rings $\mathbb{A}/\mathfrak{m}^{2^i}$; instead, as per the convention in the introduction, we assume that each coefficient reduction modulo $\mathfrak{m}^{2^i}$ takes time T_{2^i}, for a total of $td^2 T_{2^i}$ for each $i = 1, \dots, k$. This concludes the proof of our main theorem. When we work with the punctual Gröbner cell, we saw in Proposition 4.4.4 that only δm_s coefficients of each input polynomial are needed, whence $t\delta m_s T_{2^i}$ steps for coefficient reduction, for all indices i.

Remark 4.7.1. If one wishes to work only with Gröbner bases as input and output, it is straightforward to design algorithms called LiftOneStepGroebnerBasis (and LiftOneStepPunctualGroebnerBasis), that take $f'_1, \dots, f'_t$ and $\mathcal{G}$ mod $\mathfrak{m}^\kappa$ as input and return $\mathcal{G}$ mod $\mathfrak{m}^{2\kappa}$. It suffices to call Algorithm ParametersFromReducedBasis when entering the procedure, then Algorithm LiftOneStep, and finally Algorithm ReducedBasisFromParameters before exiting (or their punctual variants). This does not affect asymptotic runtimes, but is not useful in the context of our main theorem.

Remark 4.7.2. When $\mathfrak{m}$ is principal, we can slightly improve of the lifting procedure by using either divide-and-conquer techniques (folklore) or relaxed algorithms [17, Section 4] to solve the linear system that gives $\epsilon_1, \dots, \epsilon_N$. The downside is that the runtime is not written in terms of operations in $\mathbb{A}_{2\kappa}$ anymore. Instead, we give runtimes for the common cases $\mathbb{A} = \mathbb{Z}$ and $\mathfrak{m} = \langle p \rangle$, and $\mathbb{A} = \Bbbk[t]$ and $\mathfrak{m} = \langle t - \tau \rangle$:

- In the former case, solving the system uses $O\tilde{}(t\delta^\omega \log(p))$ bit operations, for a one-time computation (matrix inversion) done modulo p, and $O\tilde{}(\delta^2 \kappa \log(p))$ for subsequently solving the system modulo $p^{2\kappa}$.

- In the latter case, the one time computation takes $O\tilde{}(t\delta^\omega)$ operations in $\Bbbk$, after which linear system solving takes $O\tilde{}(\delta^2\kappa)$ operations in $\Bbbk$.

To wit, each operation in $\mathbb{A}_{2\kappa}$, as reported in Proposition 4.7.2, takes $O\tilde{}(\kappa \log(p))$ bit operations in the former case, and $O\tilde{}(\kappa \log(p))$ operations in the latter. The net effect is that in both case, the cost of solving the linear system can be neglected (up to the one-time computation we perform at the beginning).

❧ **Example 4.7.1**

We show one step of the algorithm for our running example (Example 4.1.1), focusing on the punctual Gröbner parameters. Our input is the polynomials f_1, f_2 as in the introduction, together with the Gröbner basis of the $\langle x, y\rangle$-primary component of $\langle f_1 \text{ rem } p, f_2 \text{ rem } p\rangle$, with $p = 11$; namely:

$$\left|\begin{array}{l} y^4 + 2xy + 7x^2, \\ xy^3 + 5x^3, \\ x^2y + 9x^3, \\ x^4. \end{array}\right.$$

We deduce the punctual Gröbner parameters modulo 11, $\alpha = (0, 2, 2, 5, 9) \in \mathbb{Z}/11\mathbb{Z}^5$ (recall that $N' = 5$ here). Following the algorithm, we set $(\ell_1, \dots, \ell_5) = (\Lambda_1, 2 + \Lambda_2, 2 + \Lambda_3, 5 + \Lambda_4, 9 + \Lambda_5)$ and we compute the corresponding punctual Gröbner basis, with coefficients truncated modulo 11^2 and $\langle \Lambda_1, \dots, \Lambda_N\rangle^2$. We obtain the polynomials written $\mathcal{G}^*$ in the pseudo-code:

$$\left|\begin{array}{l} y^4 + \Lambda_1 xy^2 + (\Lambda_2 + 2)xy + (40\Lambda_1 + \Lambda_3 + 103\Lambda_4 + 111\Lambda_5 + 33)x^3 + (9\Lambda_2 + 2\Lambda_5 + 18)x^2, \\ xy^3 + (\Lambda_4 + 5)x^3, \\ x^2y + (\Lambda_5 + 9)x^3, \\ x^4. \end{array}\right.$$

Reducing f_1 and f_2 modulo $\mathcal{G}^*$ (with calculations done modulo 11^2 and $\langle \Lambda_1, \dots, \Lambda_N\rangle^2$), and keeping coefficients, we obtain the linear equations $\mathscr{R}$ (we only show the non-zero ones)

$$14\Lambda_1 = 14\Lambda_2 + 11 = 76\Lambda_1 + 14\Lambda_3 + 111\Lambda_4 + 102\Lambda_5 + 99 = 5\Lambda_2 + 28\Lambda_5 + 44 = 103\Lambda_4 + 10\Lambda_5 = 0.$$

They admit the following unique solution modulo 11^2:

$$\epsilon_1 = 0,\ \epsilon_2 = 77,\ \epsilon_3 = 110,\ \epsilon_4 = 88,\ \epsilon_5 = 110;$$

as expected, all ϵ_i vanish modulo 11. From this, α is updated to take the value $\alpha+\epsilon = [0, 79, 112, 93, 119]$ modulo 11^2. One can verify that this coincides modulo 11^2 with the values given in 4.6.0.4.

4.8 Conclusion

A natural question is whether our approach can be used for ideals in more than two variables. As of now, several ingredients are missing: the known structure results are not as complete as Lazard's [123], and there is no known explicit description of Gröbner cells. Algorithmically, the key operation (reduction modulo an n-variate lexicographic Gröbner basis) seems to be a challenging problem in itself.

As already mentioned in the introduction, using our results in order to recover $\mathcal{G}$ itself, rather than $\mathcal{G}$ rem $\mathfrak{m}^K$, including in particular the quantification of bad maximal ideals $\mathfrak{m}$, is the subject of future work. Beyond this, the main algorithmic improvement we would like to achieve is reducing the overall cost so that it matches that of [111], in cases where both approaches are applicable.

Chapter 5

$\mathfrak{m}$-adic algorithm for bivariate Gröbner bases

Overview of this Chapter Let $\mathbb{A}$ a domain and $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal. We present an $\mathfrak{m}$-adic algorithm to recover the lexicographic Gröbner basis $\mathcal{G}$ of a zero-dimensional ideal $\langle \mathcal{F} \rangle \subseteq \mathbb{K}[x, y]$ with $\mathcal{F} \subseteq \mathbb{A}[x, y]$, where $\mathbb{A}$ is contained in a field $\mathbb{K}$ endowed with a valuation, and for which there exists a fraction reconstruction algorithm. We observe that previous results of Lazard's that use Hermite normal forms to compute Gröbner bases for ideals with two generators can be generalised to a set of $t \in \mathbb{Z}^+$ generators. We use this result to obtain a bound on the height of the coefficients of $\mathcal{G}$, and to control the probability of choosing a *good* maximal ideal $\mathfrak{m} \subseteq \mathbb{A}$ to build the $\mathfrak{m}$-adic expansion of $\mathcal{G}$. We complete the cost analysis when $\mathbb{A} = \mathbb{Z}$ and we obtain a complexity that is less than cubic in terms of the dimension of $\mathbb{Q}[x, y]/\langle \mathcal{G} \rangle$ and softly linear in the height of its coefficients.

To be published (in part): É. Schost, and C. St-Pierre. p-adic algorithm for bivariate Gröbner bases. In *ISSAC'23*, pages (to be determined). ACM Press, 2023 doi: 10.1145/3597066.3597086

5.1 Introduction

There already exists a rich literature dedicated to the solutions of systems of polynomial equations in two variables [84, 67, 56, 4, 141, 14, 66, 27, 111, 25, 102, 124, 103, 26, 54, 48], due in part to their numerous applications in real algebraic geometry and computer-aided design.

Our focus in this chapter is on the complexity of computing the lexicographic Gröbner basis of a zero-dimensional ideal in $\mathbb{K}[x, y]$ with a generating set in $\mathbb{A}[x, y]$ for $\mathbb{A} \subseteq \mathbb{K}$ a domain, specifically by means of $\mathfrak{m}$-adic techniques, where $\mathfrak{m} \subseteq \mathbb{A}$ is a maximal ideal, based on Newton iterations. We give a general algorithm and a complete example of the complexity analysis when $\mathbb{A} = \mathbb{Z}$ and $\mathbb{A} = k[\boldsymbol{t}]$ for k a large field. An important aspect of this work [this chapter] is to give size bounds for such a Gröbner basis, as well as bounds on the number of maximal ideals of ill $\mathfrak{m}$-adic expansion.

Over $\mathbb{Z}$, p-adic techniques have been considered in the context of Gröbner basis computations (in an arbitrary number of variables) for decades. In 1983 and 1984, Ebert and Trinks addressed the question of modular algorithms for Gröbner bases [59, 153], specifically for systems without multiple roots; these techniques were used as well in the geometric resolution algorithm [82, 81, 79, 83]. The absence of multiple roots allows for simple and efficient algorithms; for arbitrary inputs, the question is more involved.

Winkler gave the first p-adic algorithm to construct a Gröbner basis [155] that applies to general inputs; Pauer refined the discussion of good prime numbers [135], and Arnold revisited, and simplified, these previous constructions in [6]. No complexity analysis was provided; these p-adic algorithms remain complex (they lift not only the Gröbner basis but also the transformation matrix that turns the input system into its Gröbner basis), and to our knowledge, achieve linear convergence. It is desirable, although never shown so far, that $\mathfrak{m}$-adic algorithm with a quadratic convergence rate, like it is naturally the case in Newton iteration, could be used to find Gröbner bases. This is fulfilled in this document in the case of bivariate zero-dimensional ideal, *i.e.* intersections of plane curves.

In Chapter 4 [143], we presented a form of Newton iteration specifically tailored to lexicographic Gröbner bases in two variables, with no assumptions on the input. It crucially rests on results due to Conca and Valla [40], who gave an explicit parametrization of bivariate ideals with a given initial ideal: our lifting algorithm works specifically with the parameters introduced by Conca *&* Valla. Our contribution in this paper is to build on [Chapter 4][143] to give a complete $\mathfrak{m}$-adic algorithm: we quantify bad maximal ideals, show how to initialize the lifting process, give bounds on the size of the output, and analyze the cost of the whole algorithm over $\mathbb{Q}$ and $k(\boldsymbol{t})$.

The following theorem gives the results over $\mathbb{Z}$, where the probability of success and the number of input polynomials are kept constant (the more precise version is given in the last section). In what follows, the *height* of a nonzero integer u is $\log(|u|)$; if $\mathcal{G}$ is a family of polynomials in $\mathbb{Q}[x, y]$, we define $\deg(\mathcal{G}) = \dim_{\mathbb{Q}} \mathbb{Q}[x, y]/\langle \mathcal{G} \rangle$ and let $H(\mathcal{G})$ be the maximum height of the numerators and denominators of its coefficients.

Theorem 5.1.1. *Let $\mathcal{F} = (f_1, \dots, f_t)$ be in $\mathbb{Z}[x, y]$, with degree at most d, height at most h, and with finitely many common solutions in $\mathbb{C}^2$. Let $\mathcal{G}$ be the lexicographic Gröbner basis of $I = \langle \mathcal{F} \rangle \subseteq \mathbb{Q}[x, y]$ for the order $x \prec y$ and write*

$$s = |\mathcal{G}|, \quad \delta = \deg(\mathcal{G}), \quad b = H(\mathcal{G}).$$

For $P > 0$, assuming $P \in O(1)$ and $t \in O(1)$, there exists an algorithm that computes $\mathcal{G}$ with probability of success at least $1 - 1/2^P$ using a number of bit operations softly linear in

$$d^2 h + (d^{\omega+1} + \delta^{\omega}) + (d^2\delta + d\delta^2 + s^2\delta^2)b.$$

With the notation in the theorem, the bitsize of the input is linear in d^2h, and that of the output is linear in $s\delta b$.

If all points of $V(I)$ have multiplicity one, *i.e.* the local structure of all points is trivial, previous forms of Newton iteration achieves better runtimes, softly linear in the output size [83, 142, 49, 111] (but instead of a Gröbner basis, they compute a *triangular decomposition* of $V(I)$, or change coordinates). Hence, it makes sense to consider applying our techniques only to multiple solutions.

This is what motivates our second result, where we compute the Gröbner basis of the $\langle x, y \rangle$-primary component J of I. While this remains to be done, a natural extension of this result is to combine it with those in Chapter 3 [94], which shows how to put an arbitrary primary component of I in correspondence with the $\langle x, y \rangle$-primary component of a related ideal in $\mathbb{K}[x, y]$, for a finite extension $\mathbb{K}$ of $\mathbb{Q}$.

Theorem 5.1.2. *Let $\mathcal{F} = (f_1, \dots, f_t)$ be in $\mathbb{Z}[x, y]$, with degree at most d, height at most h, and with finitely many common solutions in $\mathbb{C}^2$. Let $\mathcal{G}^0$ be the lexicographic Gröbner basis of the $\langle x, y \rangle$-primary component of $I = \langle \mathcal{F} \rangle \subseteq \mathbb{Q}[x, y]$ for the order $x \prec y$ and write*

$$r = |\mathcal{G}^0|, \quad \eta = \deg(\mathcal{G}^0), \quad c = H(\mathcal{G}^0).$$

For $P > 0$, assuming $P \in O(1)$ and $t \in O(1)$, there is an algorithm that computes $\mathcal{G}^0$ with probability of success at least $1 - 1/2^P$ using a number of bit operations softly linear in

$$t(d^2h + (d^{\omega}\eta + \eta^{\omega})\log(h) + \eta^2 c).$$

5.1.1 Leitfaden

Inspired by Lazard [109], we prove in Section 5.2 that the Hermite Normal form of an "extended Sylvester matrix" built from $f_1, \ldots, f_t$ gives the coefficients of what we will call a *detaching basis* of the ideal I they generate. We also present a variant of this result, where replacing Hermite normal form by Howell normal form yields a Gröbner basis of a localization of I.

We use these results in two manners: to compute the initial Gröbner basis in $\mathbb{A}/\mathfrak{m}$ for $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal, prior to entering Newton iteration, and to obtain a height bounds for the output (over $\mathbb{K}$) and quantify bad choices of bad maximal ideal $\mathfrak{m}$. Revisiting the proof from Galligo [72], Bayer-Stillman[11] and Pardue [134], we also give a constructive proof that the initial term ideals of zero-dimensional ideals in generic coordinates are Borel-fixed to characterize the Zariski open set. The result could be extrapolated from [144], since lexicographic ordering can be obtained with a weighted ordering. However, it does not lead usable description of a hyperplane from which we may deduce the probability that a change of coordinate makes an ideal Borel-fixed. Under the assumption of genericity for the coordinates, it implies that $\min\{i : y^i \in (I)\}$ a better bound than the degree of the I for where I is a zero-dimensional ideal to rewrite complexity of [Chapter 4] [143]. The algorithms underlying the theorems above are in Section 5.5.

5.2 Lexicographic Gröbner bases via matrix normal forms

In this section, we assume $I = \langle f_1, \ldots, f_t \rangle \subset \mathbb{K}[x, y]$, for $t \geq 2$, and we show how to derive the lexicographic Gröbner basis of I, or its primary component at the origin, from either Hermite or Howell normal forms of matrices over $\mathbb{K}[x]$, for an arbitrary field $\mathbb{K}$. These results are direct extensions of previous work of Lazard's [109], who used Hermite forms in the case $t = 2$.

In what follows, for a subset $S \subset \mathbb{K}[x, y]$ and $n \geq 0$, we let $S_{<(.,n)}$ be the subset of all f in S with $\deg_y(f) < n$; notation such as $S_{\leq(.,n)}$ is defined similarly. In particular, if S is an ideal of $\mathbb{K}[x, y]$, $S_{<(.,n)}$ is a free $\mathbb{K}[x]$-module of rank at most n. For $S = \mathbb{K}[x, y]$ itself, $\mathbb{K}[x, y]_{<(.,n)}$ is a free $\mathbb{K}[x]$-module of rank n, equal to $\bigoplus_{0 \leq i < n} \mathbb{K}[x] y^i$.

For such an n, we also let π_n denote the $\mathbb{K}[x]$-module isomorphism $\mathbb{K}[x, y]_{<(.,n)} \to \mathbb{K}[x]^n$, which maps $f_0 + \cdots + f_{n-1} y^{n-1}$ to the vector $[f_{n-1} \cdots f_0]^\top$.

5.2.1 Detaching bases

Let I be an ideal in $\mathbb{K}[x,y]$ and let $\mathcal{G} = (g_0, \dots, g_s)$ be its reduced minimal Gröbner basis for the lexicographic order induced by $y \succ x$, listed in decreasing order; we write $n_i = \deg_y(g_i)$ for all i (so these exponents are decreasing). We define polynomials $A_0, A_1, \dots$ as follows:

- for $0 \le i < n_s$, $A_i = 0$,
- if there exists k in $\{0, \dots, s\}$ such that $n_k = i$, $A_i = g_k$
- otherwise, A_i is obtained by starting from yA_{i-1}, and reducing all its terms of y-degree less than i by $\mathcal{G}$.

For example, if I has a Gröbner basis of the form $(y - f(x), g(x))$, the polynomials A_i are given by $A_0 = g$, $A_1 = y - f$ and for $i \ge 2$, $A_i = y^i - (f^i \bmod g)$. See for instance [9] for a previous discussion of such bases.

Lemma 5.2.1. *For $i \ge n_s$, $\deg_y(A_i) = i$.*

Proof. This is true for i of the form n_k. For i in $n_k, \dots, n_{k-1} - 1$, we proceed by induction, with the remark above establishing the base case (for $k = 0$, we consider all $i \ge n_0$). Assume $\deg_y(A_{i-1}) = i - 1$, so that $\deg_y(yA_{i-1}) = i$. Because we use the lexicographic order $x \prec y$, the reduction of the terms of y-degree less than i in yA_{i-1} does not introduce terms of y-degree i or more. □

Lemma 5.2.2. *For $n \ge n_s$, the $\mathbb{K}[x]$-module $I_{\le(.,n)}$ is free of rank $n - n_s + 1$, with basis $A_{n_s}, \dots, A_n$.*

Proof. The polynomials $A_{n_s}, \dots, A_n$ are all nonzero, with pairwise distinct y-degrees, so they are $\mathbb{K}[x]$-linearly independent. Visibly, they all belong to $I_{\le(.,n)}$, so it remains to prove that they generate $I_{\le(.,n)}$, as a $\mathbb{K}[x]$-module.

This is done by induction on $n \ge n_s$. Take f in $I_{\le(.,n)}$, and write it as $f = f_n y^n + g$, with f_n in $\mathbb{K}[x]$ and g in $K[x,y]_{\le(.,n-1)}$. Let $h_n \in \mathbb{K}[x]$ be the polynomial coefficient of y^n in A_n, so that $A_n = h_n y^n + B_n$, with B_n in $\mathbb{K}[x,y]_{\le(.,n-1)}$. Write the Euclidean division $f_n = qh_n + r$ in $\mathbb{K}[x]$, with $\deg_x(r) < \deg_x(h_n)$, and rewrite f as

$$\begin{aligned} f &= (qh_n + r)y^n + g \\ &= qh_n y^n + ry^n + g \\ &= qA_n - qB_n + ry^n + g. \end{aligned}$$

The polynomial $-qB_n + ry^n + g$ is in I, so its normal form modulo $\mathcal{G}$ is zero. The terms $-qB_n + g$ have y-degree less than n, so their normal form has y-degree less than n as well; since ry^n is already reduced modulo $\mathcal{G}$, it must be zero.

It follows that $f = qA_n + g - qB_n$, with $g - qB_n$ in $I_{\leq(.,n-1)}$. If $n = n_s$, this latter polynomial must vanish; this proves the base case of our induction. Else, by induction assumption, it is a $\mathbb{K}[x]$-linear combination of $A_{n_s}, \ldots, A_{n-1}$; this finishes the proof. □

For $n \geq n_0$, the *detaching basis* of I in degree n is the sequence $(A_{n_s}, \ldots, A_n)$. Because we take $n \geq n_0$, this is (in general) a non-minimal Gröbner basis of I, and we can recover $\mathcal{G}$ from it by discarding redundant entries (that is, all polynomials whose leading term is a multiple of another leading term).

5.2.2 Using Hermite normal forms

Given $\mathcal{F} = (f_1, \ldots, f_t)$ in $\mathbb{K}[x, y]$, we prove that the Hermite normal form of a certain Sylvester-like matrix associated to them gives a lexicographic detaching basis of the ideal I they generate. In [109], Lazard covered the case $t = 2$, under an assumption on the leading coefficients (in y) of the f_i's.

We extend his work (in a direct manner) to situations where such assumptions do not hold. First, to polynomials $\mathcal{F} = (f_1, \ldots, f_t)$ in $\mathbb{K}[x, y]$, we associate an integer $\Delta(\mathcal{F})$, defined as follows.

❧ **Definition 28.** *Let $\mathcal{F} = (f_1, \ldots, f_t)$ be in $\mathbb{K}[x, y]$ and let $(A_{n_s}, \ldots, A_{n_0})$ be their detaching basis in degree n_0, with n_0 and n_s the maximal, resp. minimal y-degree of the polynomials in the lexicographic Gröbner basis of $\langle f_1, \ldots, f_t\rangle$, for the order $x \prec y$.*

We let $\Delta(\mathcal{F})$ be the minimal integer Δ such that for $i = n_s, \ldots, n_0$, there exist $w_{i,1}, \ldots, w_{i,t}$ in $\mathbb{K}[x, y]^t$, all of y-degree less than Δ, and such that $A_i = w_{i,1} f_1 + \cdots + w_{i,t} f_t$.

The following proposition gives the basic application we will make of this integer, allowing us to extract a detaching basis from a Hermite form computation. Our convention for Hermite normal forms (here, for matrices over $\mathbb{K}[x]$) is the following: we use *column* operations, with Hermite normal forms being lower triangular. The first nonzero entry in a nonzero column is called its *pivot*, its index being called the *pivot index*. By convention, pivots in nonzero columns of a matrix in Hermite form are monic in x.

Proposition 5.2.1. *Let $\mathcal{F} = (f_1, \ldots, f_t)$ be in $\mathbb{K}[x, y]$, for $t \geq 2$, of y-degree at most d_y, and assume that they generate an ideal $I = \langle f_1, \ldots, f_t\rangle$ of dimension zero. For $i = 1, \ldots, t$, write $f_i = f_{i,0} + \cdots + f_{i,d_y} y^{d_y}$, with all $f_{i,j}$ in $\mathbb{K}[x]$.*

For $D \geq \Delta(\mathcal{F})$, let $c_1, \dots, c_K$ be the nonzero columns of the Hermite normal form $\boldsymbol{H}$ of $\boldsymbol{S} = [\boldsymbol{S}_1 \cdots \boldsymbol{S}_t] \in \mathbb{K}[x]^{(d_y+D)\times tD}$, where

$$\boldsymbol{S}_i = \begin{bmatrix} f_{i,d_y} & & \\ \vdots & \ddots & \\ f_{i,0} & & f_{i,d_y} \\ & \ddots & \vdots \\ & & f_{i,0} \end{bmatrix} \in \mathbb{K}[x]^{(d_y+D)\times D}.$$

Then, there exists $K' \leq K$ such that $\pi^{-1}_{d_y+D}(c_{K'})$ is monic in y; with K' the largest such integer, $\pi^{-1}_{d_y+D}(c_K), \dots, \pi^{-1}_{d_y+D}(c_{K'})$ is a detaching basis of I.

In particular, while we do not know the y-degrees n_i of the elements in the Gröbner basis of I, as long as $D \geq \Delta(\mathcal{F})$, it is enough to consider the last nonzero columns of $\boldsymbol{H}$, stopping when we find (through $\pi^{-1}_{d_y+D}$) a polynomial that is monic in y. Remark also that we do not assume that the polynomials f_i have y-degree exactly d_y.

Proof. Let $D \geq \Delta(\mathcal{F})$ be as in the proposition. Let us index the columns of each block $\boldsymbol{S}_i$ by $y^{D-1}, \dots, y, 1$, and its rows by $y^{d_y+D-1}, \dots, y, 1$. Then, $\boldsymbol{S}_i$ is the matrix of the map $\mathbb{K}[x,y]_{<(.,D)} \to \mathbb{K}[x,y]_{<(.,d_y+D)}$ given by $w_i \mapsto w_i f_i$. The matrix $\boldsymbol{S}$ itself maps a vector $(w_1, \dots, w_t)$, with all entries of y-degree less than D, to $\sum_{i=1}^t w_i f_i \in I_{<(.,d_y+D)}$.

Let $\mathcal{G} = (g_0, \dots, g_s)$ be the lexicographic Gröbner basis of $I = \langle f_1, \dots, f_t \rangle$, listed in decreasing order, with $\deg_y(g_i) = n_i$ for all i. Since we assume that I has dimension zero, we have $n_s = 0$, and g_0 is monic in y.

Let $A_0, \dots, A_{n_0}$ be the detaching basis of I in degree n_0. We denote by $c_1, \dots, c_K$ the nonzero columns of the Hermite form $\boldsymbol{H}$ of $\boldsymbol{S}$, and we let $H_i = \pi^{-1}_{d_y+D}(c_i)$, for $i = 0, \dots, n_0$. We will prove that $A_i = H_{K-i}$ for $i = 0, \dots, n_0$. Since g_0 is the only polynomial in $A_0, \dots, A_{n_0}$ which is monic in y, this will establish the proposition, with $K' = K - n_0$.

Since both A_i and H_{K-i} are in I, to prove that they are equal, it is enough to prove that for all i, $A_i - H_{K-i}$ is reduced with respect to the Gröbner basis $\mathcal{G}$ of I.

Because $D \geq \Delta(\mathcal{F})$, we deduce that $A_0, \dots, A_{n_0}$ are in the column span of $\boldsymbol{S}$. Since they have respective y-degrees $0, \dots, n_0$, we see that $\deg_y(H_{K-i}) = \deg_y(A_i) = i$ for all $i = 0, \dots, n_0$. In addition, for all such i, we can write $A_i = \sum_{j=0}^i a_{i,j} H_{K-j}$, for some $a_{i,j}$ in $\mathbb{K}[x]$.

On the other hand, Lemma 5.2.2 shows that for the same index i, we can write $H_{K-i} = \sum_{j=0}^i b_{i,j} A_j$, for some $b_{i,j}$ in $\mathbb{K}[x]$. Because both A_i and H_{K-i} have leading

y-coefficients that are monic in x, it follows that $b_{i,i} = a_{i,i} = 1$ for all i. This proves that A_i and H_{K-i} have the same coefficient of y-degree i (call it $M_i \in \mathbb{K}[x]$), and thus that $A_i - H_{K-i}$ has y-degree less than i.

By definition of a detaching basis, all terms of y-degree less than i in A_i are reduced with respect to $\mathcal{G}$. On the other hand, by the property of Hermite forms, for $j < i$, the coefficient of y-degree j in H_{K-i} is reduced with respect to M_j. Since we saw that M_j is also the coefficient of y^j in A_j, this proves that all terms of y-degree less than i in H_{K-i} are reduced with respect to $A_0, \ldots, A_{i-1}$, and thus with respect to $\mathcal{G}$. Altogether, $A_i - H_{K-i}$ itself is reduced with respect to $\mathcal{G}$, which is what we set out to prove. □

We call HermiteGroebnerBasis$(\mathcal{F}, D)$ a procedure that takes as input $\mathcal{F} = (f_1, \ldots, f_t)$ and D, and returns the lexicographic Gröbner basis of $I = \langle f_1, \ldots, f_t \rangle$ obtained by computing the Hermite normal form of $\boldsymbol{S}$ as above, extracting the Gröbner basis of I from its detaching basis. Here, we take for d_y the maximum degree of the f_i's, and we assume that we have $D \geq \Delta(\mathcal{F})$ and $D \geq d_y$.

The assumption that the ideal I has dimension zero implies that it contains a non-zero polynomial in $\mathbb{K}[x]$; as a result, its detaching basis has entries of y-degrees $0, 1, \ldots,$ so that the Hermite form of $\boldsymbol{S}$ is lower triangular with $d_y + D$ non-zero diagonal entries. In other words, $\boldsymbol{S}$ has rank $d_y + D$ (seen as a matrix over $\mathbb{K}(x)$).

If $t = 2$ and $D = d_y$, this matrix is square, but in general, it may have more columns than rows (recall that we assume $D \geq d_y$). Using the algorithm of [106], we can permute the columns of $\boldsymbol{S}$ to find a $(d_y + D) \times tD$ matrix $\boldsymbol{S}'$ whose leading $(d_y + D) \times (d_y + D)$ minor is nonzero; this takes $O\tilde{~}(tD^\omega d)$ operations in $\mathbb{K}$, with d the maximum degree of the f_i's. Let us define the $tD \times tD$ square matrix

$$\boldsymbol{S}^{\mathrm{sq}} = \begin{bmatrix} & \boldsymbol{S}' \\ \mathbf{0}_{(t-1)D-d_y, d_y+D} & \boldsymbol{I}_{(t-1)D-d_y,(t-1)D-d_y} \end{bmatrix} \qquad (5.2.0.1)$$

together with its Hermite form $\boldsymbol{H}^{\mathrm{sq}}$; the first $d_y + D$ rows of it give us the Hermite form $\boldsymbol{H}$ of $\boldsymbol{S}$. The Hermite form of $\boldsymbol{S}^{\mathrm{sq}}$ is computed in $O\tilde{~}(t^\omega D^\omega d)$ operations in $\mathbb{K}$ [107]. This gives the overall cost of computing the lexicographic Gröbner basis of I, assuming an upper bound on $\Delta(\mathcal{F})$ is known. To our knowledge, not much exists in the literature on complete cost analysis for bivariate ideals, apart from BurchBerger's algorithm is bounded by $\frac{3}{2}(t + 2(d+2)^2)^4$ operations in the field [34].

The following proposition gives various bounds on $\Delta(\mathcal{F})$, whose strength depends on the assumptions we make on $\mathcal{F}$. The first one is a direct extension of Lazard's [109, Lemma 7], and is linear in the y-degree of the input. The others are based on results from [104, 55], which involve total degree considerations.

Proposition 5.2.2. *Let $\mathcal{F} = (f_1, \dots, f_t)$ be in $\mathbb{K}[x, y]$ of degree at most $d \geq 1$, and y-degree at most d_y, and let $I = \langle f_1, \dots, f_t \rangle \subset \mathbb{K}[x, y]$. Set $d' = \max(d, 3)$. Then:*

- *if there exists i in $\{1, \dots, t\}$ such that the coefficient of y^d in f_i is a nonzero constant, $\Delta(\mathcal{F}) \leq \Delta_1(d_y) = d_y$*
- *if $t = 2$ and I has finitely many zeros over $\overline{\mathbb{K}}$, $\Delta(\mathcal{F}) \leq \Delta_2(d) = 2d'^2 + d' \in O(d^2)$*
- *if I has finitely many zeros over $\overline{\mathbb{K}}$, $\Delta(\mathcal{F}) \leq \Delta_3(d) = 16d'^4 + 2d'^2 + 2d' \in O(d^4)$*

First item. In what follows, without loss of generality, we assume that the coefficient of y^{d_y} in f_t is 1. We prove a slightly more general claim: *any polynomial* f in $I_{<(.,2d_y)}$ can be written as $f = w_1 f_1 + \dots + w_t f_t$, with all w_i in $\mathbb{K}[x, y]_{<(.,d_y)}$. This is enough to conclude, since (with the notation used in the definition of Δ) all entries $A_{n_s}, \dots, A_{n_0}$ in the detaching basis of I in degree n_0 have y-degree at most $d_y \leq 2d_y - 1$ (this is because we use a lexicographic order with $x \prec y$).

Let thus f be given in $I_{<(.,2d_y)}$. There exists at least one family $w = (w_1, \dots, w_t)$ in $\mathbb{K}[x, y]$ such that

$$f = \sum_{i=1}^{t} w_i f_i, \tag{5.2.0.2}$$

since f is in I. For such a family w, we define $\mathcal{S}_w = \{i \mid \deg_y(w_i) \geq d_y\}$. For any w such that $\mathcal{S}_w$ is not empty, we further set $\nu_w = \min(\mathcal{S}_w) \in \{1, \dots, t\}$, and we let ν be the *maximal* value of these ν_w's. To see that ν is well-defined, note that there is a vector w for which $\mathcal{S}_w$ is not empty (we can replace (w_{t-1}, w_t) by $(w_{t-1} + g f_t, w_t - g f_{t-1})$ for any g in $\mathbb{K}[x, y]$).

Let w be such that $\nu = \nu_w$. We claim that $\mathcal{S}_w \neq \{t\}$: otherwise we would have $\deg_y(w_t f_t) \geq 2d_y$, while $\deg_y(w_i f_i) < 2d_y$ for all other i's; this would contradict the assumption $\deg_y(f) < 2d_y$. This shows that $\nu < t$.

Let us further refine our choice of w, by taking it such that, among all those vectors for which $\mathcal{S}_w$ is not empty and $\nu_w = \nu$, the y-degree of w_ν is minimal. Let us then write $e = \deg_y(w_\nu)$ (so that $e \geq d_y$) and let $c \in \mathbb{K}[x]$ be the coefficient of y^e in w_ν. We can use it to rewrite f as

$$f = \sum_{i=1}^{t} w_i f_i + c y^{e-d_y} f_\nu f_t - c y^{e-d_y} f_t f_\nu.$$

If we set

$$w'_i = \begin{cases} w_\nu - c y^{e-d_y} f_t & \text{when } i = \nu; \\ w_t + c y^{e-d_y} f_\nu & \text{when } i = t; \\ w_i & \text{otherwise,} \end{cases}$$

we still have

$$f = \sum_{i=1}^{t} w'_i f_i.$$

By construction, $\deg_y(w'_i) = \deg_y(w_i) < d_y$ for all $i < \nu$, so none of $1, \ldots, \nu - 1$ is in $\mathcal{S}_{w'}$. If ν is in $\mathcal{S}_{w'}$, then the inequality $\deg_y(w'_\nu) < \deg_y(w_\nu)$ contradicts the choice of w, so that ν is not in $\mathcal{S}_{w'}$. This shows that $\mathcal{S}_{w'}$ is empty, since otherwise its minimum element would be greater than ν. □

For the second and third items, we use results from [55], for which we need total degree bounds on the input polynomials $\mathcal{F} = (f_1, \ldots, f_t)$ and the elements $A_0, \ldots, A_{n_0}$ in the detaching basis (here, $n_s = 0$, since I having finitely many solutions implies that it contains a nonzero polynomial in $\mathbb{K}[x]$). For the inputs f_i, we have the degree bound $\deg(f_i) \le d \le d'$. For the A_i's, we have the bounds $\deg_x(A_i) \le d^2$ (by Bézout's theorem) and $\deg_y(A_i) \le d$, for $i \le n_0$, so their total degree is at most $D = d'^2 + d'$.

Second item. When $t = 2$ and I has dimension zero (that is, has a finite, nonzero number of solutions in $\overline{\mathbb{K}}$), f_1, f_2 are in complete intersection, so that we have $A_i = w_{i,1}f_1 + w_{i,2}f_2$, with $\deg_y(w_{i,j}) \le D + d'^2$ for all i, j, by Theorem 5.1 in [55]. Overall, the resulting degree bound is $2d'^2 + d'$.

If we assume that $I = \mathbb{K}[x, y]$, we know that there are g_1, g_2 in $\mathbb{K}[x, y]$ such that $g_1 f_1 + g_2 f_2 = 1$, with $\deg(g_i) \le d'^2$ [104]. Multiplying this by A_j, for $j \le n_0$, we obtain the expression $(g_1 A_j) f_1 + (g_2 A_j) f_2 = A_j$, with $\deg_y(g_i A_j) \le d'^2 + d$ in this case. □

Third item. We apply Corollary 3.4 from [55]. It gives an upper bound on the total degree (and thus y-degree) of the coefficients in a membership equality $A_i = w_{i,1}f_1 + \cdots + w_{i,t}f_t$, showing that we can take $\deg_y(w_{i,j}) \le D + 16d'^4 + d'^2 + d'$ for all i, j. □

5.2.3 Using the Howell form

We now investigate how using another matrix normal form, the *Howell* form [93], yields information about certain primary components of an ideal I as above.

Howell forms are defined for matrices with entries in a principal ideal ring $\mathbb{A}$; below, we will take $\mathbb{A} = \mathbb{K}[x]/x^k$, for an integer k. As for the Hermite form, we consider column operations; then, an $n \times m$ matrix $\boldsymbol{H}$ over $\mathbb{A} = \mathbb{K}[x]/x^k$ is in Howell normal form if the following (taken from [150, Chapter 4]) hold:

1. let $r \leq m$ be the number of nonzero columns in $\boldsymbol{H}$; then these nonzero columns have indices $1, \ldots, r$
2. $\boldsymbol{H}$ is in lower echelon form: for $i = 1, \ldots, r$, let $j_i \in \{1, \ldots, n\}$ be the index of the first nonzero entry in the ith column; then, $j_1 < \cdots < j_r$
3. all pivots $H_{j_i,i}$, for $i = 1, \ldots, r$, are of the form x^{c_i}
4. for $i = 1, \ldots, r$ and $k = 1, \ldots, i-1$, $H_{j_i,k}$ is reduced modulo $H_{j_i,i}$
5. for $i = 0, \ldots, r$, any column in the column span of $\boldsymbol{H}$ with at least j_i leading zeros is an $\mathbb{A}$-linear combination of columns of indices $i+1, \ldots, r$ (here, we set $j_0 = 0$)

For any matrix $\boldsymbol{M}$ in $\mathbb{A}^{n \times m}$, there is a unique $\boldsymbol{H}$ in Howell normal form in $\mathbb{A}^{n \times m}$, and a not necessarily unique invertible matrix $\boldsymbol{U}$ in $\mathbb{A}^{m \times m}$ such that $\boldsymbol{H} = \boldsymbol{M}\boldsymbol{U}$. The matrix $\boldsymbol{H}$ is called the Howell normal form of $\boldsymbol{M}$.

Given $f_1, \ldots, f_t$ as before, we are interested here in computing the lexicographic Gröbner basis of $J = \langle f_1, \ldots, f_t, x^k \rangle$, for a given integer k. In particular, if $(0,0)$ is in $V(f_1, \ldots, f_t)$, and no other point $(0, \beta)$ is, for $\beta \neq 0$, J is the $\langle x, y \rangle$-primary component of $I = \langle f_1, \ldots, f_t \rangle$, if k is large enough (note k is large if $k[t]$ has a large characteristic or characteristic 0).

The following proposition shows how to reduce this computation to a Howell normal form calculation. In what follows, the *canonical lift* of an element in $\mathbb{A} = \mathbb{K}[x]/x^k$ to $\mathbb{K}[x]$ is its unique preimage of degree less than k; this carries over to vectors and matrices (and in particular to the output of the Howell form computation). Contrary to what happens for Hermite forms, there is no guarantee that the polynomials extracted from the Howell form are a detaching basis, as we may be missing the first polynomial (that belongs to $\mathbb{K}[x]$) and its multiples. The proposition below restores this by considering a few extra columns, if needed.

Proposition 5.2.3. *Let $f_1, \ldots, f_t$ be in $\mathbb{K}[x,y]$, for $t \geq 2$, of y-degree at most d_y, and assume that they generate an ideal of dimension zero. Let k be a positive integer and $\mathbb{A} = \mathbb{K}[x]/x^k$.*

For $D \geq \Delta(f_1, \ldots, f_t, x^k)$, let $\mathcal{B} \in \mathbb{A}^{(d_y+D) \times tD}$ be the Howell normal form of $\bar{\boldsymbol{S}} = \boldsymbol{S} \bmod x^k$, with $\boldsymbol{S}$ as in Proposition 5.2.1, and let $\boldsymbol{B}_{\mathrm{lift}}$ be its canonical lift to $\mathbb{K}[x]^{(d_y+D) \times tD}$.

Let $h_1, \ldots, h_L$ be the nonzero columns of $\mathcal{B}_{\mathrm{lift}}$, and let $r \in \{1, \ldots, d_y + D\}$ be the pivot index of h_L. Set $L'' = L + d_y + D - r$ and, for $i = L+1, \ldots, L''$ let $h_i = [0 \ \cdots \ 0 \ x^k \ 0 \ \cdots \ 0]^\top$, with x^k at index $r + i - L \in \{r+1, \ldots, d_y + D\}$.

Then, there exists $L' \leq L$ such that $\pi_{d_y+D}^{-1}(h_{L'})$ is monic in y; with L' be the largest such integer, $\pi_{d_y+D}^{-1}(h_{L''}), \dots, \pi_{d_y+D}^{-1}(h_{L'})$ is a detaching basis of $\langle f_1, \dots, f_t, x^k \rangle$.

Proof. Let $\Gamma = (\Gamma_0, \dots, \Gamma_\sigma)$ be the lexicographic Gröbner basis of $J = \langle f_1, \dots, f_t, x^k \rangle$, listed in decreasing order, with Γ_i of y-degree ν_i for all i; since x^k is in J, $\nu_\sigma = 0$. We can then let $C_0, \dots, C_{\nu_0}$ be the detaching basis of J in degree ν_0, with $\deg_y(C_i) = i$ for all i.

We know that the first polynomials in the detaching basis are of the form $C_0 = x^\ell, C_1 = yx^\ell, \dots, C_{\nu_{\sigma-1}-1} = y^{\nu_{\sigma-1}-1}x^\ell$, for some $\ell \leq k$. If $\ell = k$, they all vanish modulo x^k, but the next polynomial $C_{\nu_{\sigma-1}}$ does not. If $\ell < k$, none of them vanishes modulo x^k. Thus, we define $\rho = \nu_{\sigma-1}$ in the former case and $\rho = 0$ in the latter.

Let further $D \geq \Delta(f_1, \dots, f_t, x^k)$ be as in the proposition. If we consider the extended Sylvester matrix $\boldsymbol{T} \in \mathbb{K}[x]^{(d_y+D)\times(t+1)D}$ built from $f_1, \dots, f_t, x^k$, the assumption on D shows that each $\pi_{d_y+D}(C_i)$ is in the column span of $\boldsymbol{T}$. For $i = 0, \dots, \nu_0$, we let v_i be the column vector $\pi_{d_y+D}(C_i) \bmod x^k \in \mathbb{A}^{d_y+D}$; the discussion in the previous paragraph shows that the nonzero vectors v_i are precisely $v_\rho, \dots, v_{\nu_0}$. By reduction modulo x^k of the membership relations above, we see that $v_\rho, \dots, v_{\nu_0}$ are in the $\mathbb{A}$-span of the columns of $\bar{\boldsymbol{S}}$.

Lazard's structure theorem for bivariate lexicographic Gröbner bases [109, Theorem 1] shows that every polynomial Γ_j in the reduced Gröbner basis of J is of the form $\Gamma_j = x^{m_j}\gamma_j$, with γ_j monic in y and $m_j \leq \ell$ (the inequality is strict, except for $j = 0$). It follows that for $i = \rho, \dots, \nu_0$, the pivot in v_i is also a power of x, at index $d_y + D - i$ (precisely, it is x^{m_j}, for j the largest integer such that $\nu_j \leq i$).

Let $\eta_1, \dots, \eta_L$ be the nonzero columns in the Howell form $\mathcal{B}$ of $\bar{\boldsymbol{S}}$. By definition of the Howell form, the former observation implies that for $i = \rho, \dots, \nu_0$, v_i is in the $\mathbb{A}$-span of those η_j's starting with at least $d_y + D - i - 1$ zeros. For such an i, since the entry at index $d_y + D - i$ in v_i is nonzero, there must exist (exactly) one η_j with pivot index $d_y + D - i$.

We now prove that the pivot in η_L is precisely at index $d_y + D - \rho$. Recall that we write $h_1, \dots, h_L$ for the canonical lifts of $\eta_1, \dots, \eta_L$ to vectors in $\mathbb{K}[x]^{d_y+D}$; in particular, the pivot index r of h_L, as defined in the proposition, is also the pivot index of η_L, so that our claim is that $r = d_y + D - \rho$.

Suppose that the pivot in η_L is at an index different from $d_y + D - \rho$. By the previous discussion, it can only lie at a larger index, say $m > d_y + D - \rho$; this may happen only if $\rho > 0$, in which case we saw that $\rho = \nu_{\sigma-1} = \deg_y(\Gamma_{\sigma-1})$ and $\Gamma_\sigma = x^k$.

Let $H_1, \dots, H_L$ be the polynomials obtained by applying $\pi_{d_y+D}^{-1}$ to $h_1, \dots, h_L$. It follows that H_L has y-degree $d_y + D - m < \rho = \deg_y(\Gamma_{\sigma-1})$, and x-degree less than $k = \deg_x(\Gamma_\sigma)$. Thus, H_L is reduced with respect to the Gröbner basis $\boldsymbol{\Gamma}$ of J. On the other hand, because η_L is in the column span of $\bar{\boldsymbol{S}}$, its canonical lift h_L is in the column space of $\boldsymbol{S}$, up to the addition of a vector with entries in $x^k\mathbb{K}[x]$. In other words, H_L is in J, so that H_L must be zero, a contradiction.

Thus, the pivot index of η_L is exactly $d_y + D - \rho$, that is, the same as that of v_ρ. Our previous discussion on the pivots in the vectors η_i then implies that for $i = \rho, \dots, \nu_0$, the pivot index of $\eta_{L+\rho-i}$ is $d_y + D - i$, that is, the same as that of v_i. This implies that

$$v_i = \sum_{j=\rho}^{i} \alpha_{i,j}\eta_{L+\rho-j}, \tag{5.2.0.3}$$

for some coefficients $\alpha_{i,j}$ in $\mathbb{A} = \mathbb{K}[x]/x^k$. On the other hand, all polynomials $H_L, \dots, H_{L+\rho-\nu_0}$ are in J (by the argument we used for H_L). By Lemma 5.2.2, we deduce that for $i = \rho, \dots, \nu_0$, $H_{L+\rho-i}$ can be written as $H_{L+\rho-i} = \sum_{j=\rho}^{i} \beta_{i,j}C_j$, for some coefficients $\beta_{i,j}$ in $\mathbb{K}[x]$. After application of π_{d_y+D} and reduction modulo x^k, this gives the equality

$$\eta_{L+\rho-i} = \sum_{j=\rho}^{i} \bar{\beta}_{i,j}v_j, \tag{5.2.0.4}$$

with $\bar{\beta}_{i,j} = \beta_{i,j} \bmod x^k$ for all i, j. We know that the pivots of both v_i and $\eta_{L+\rho-i}$ are powers of x (the latter, by the properties of the Howell form), so Eq. (5.2.0.3) and Eq. (5.2.0.4) show that the pivots in v_i and $\eta_{L+\rho-i}$ are the same, for $i = \rho, \dots, \nu_0$.

Back in $\mathbb{K}[x, y]$, we deduce that C_i and $H_{L+\rho-i}$ have the same coefficient in y^i, for $i = \rho, \dots, \nu_0$. Proceeding as in the proof of Proposition 5.2.1, we deduce that we actually have $C_i = H_{L+\rho-i}$ for $i = \rho, \dots, \nu_0$: we observe that their terms of y-degree less than i are reduced with respect to Γ; it follows that $C_i - H_{L+\rho-i}$ is both in J and reduced with respect to its lexicographic Gröbner basis, so it vanishes.

Taking $i = \nu_0$, we deduce in particular that $H_{L+\rho-\nu_0}$ is monic in y (and no H_i of larger index has this property), so the index L' defined in the proposition is $L' = L + \rho - \nu_0$; the corresponding polynomials are $C_{\nu_0}, \dots, C_\rho$.

Since we saw that $r = d_y + D - \rho$, the integer L'' in the proposition is $L'' = L + \rho$, and through $\pi^{-1}_{d_y+D}$, the columns $h_{L+1}, \dots, h_{L+\rho}$ become $y^{\rho-1}x^k, \dots, x^k$ (there is no such column if $\rho = 0$). These are precisely the polynomials $C_{\rho-1}, \dots, C_0$ that were missing if $\rho > 0$. □

We call HowellGroebnerBasis($\mathcal{F}, k, D$) a procedure that takes as input $\mathcal{F} = (f_1, \dots, f_t)$, k and D, and returns the lexicographic Gröbner basis of $\langle f_1, \dots, f_t, x^k\rangle$ obtained from the Howell form of $\bar{\boldsymbol{S}}$, taking for d_y the maximum of the degrees of $f_1, \dots, f_t$, and choosing for D the integer prescribed by Proposition 5.2.2. In this case, there is no need to make $\bar{\boldsymbol{S}}$ square: the algorithm of [150, Chapter 4] computes its Howell form using $O\tilde{~}(tD^\omega k)$ operations in $\mathbb{K}$.

The main application we will make of Howell form computation is to obtain the Gröbner basis of the $\langle x, y\rangle$-primary component of an ideal such as $I = \langle f_1, \dots, f_t\rangle$. In order to do so, we will assume that we are in "nice" coordinates, in the sense that the projection on the first factor $V(\langle \mathcal{F}\rangle) \to \overline{\mathbb{K}}$ is one-to-one.

Lemma 5.2.3. *Let $\mathcal{F} = (f_1, \dots, f_t)$ be in $\mathbb{K}[x, y]$, and suppose that the projection on the first factor $V(\langle \mathcal{F}\rangle) \to \overline{\mathbb{K}}$ is one-to-one. Let further J be the $\langle x, y\rangle$-primary component of $I = \langle f_1, \dots, f_t\rangle$, with m the smallest integer such that x^m is in J. Then:*

- *the smallest power of x in the ideal $H = \langle f_1, \dots, f_t, x^k\rangle$ is $x^{\min(m,k)}$.*
- *for $k \geq m$, $H = J$.*

Proof. First, we establish that $J = \langle f_1, \dots, f_t, x^m\rangle$. For one direction, all f_i's, as well as x^m, are in J by definition. Conversely, the assumption on $V(\langle \mathcal{F}\rangle)$ implies that we can write $\langle f_1, \dots, f_t\rangle = JJ'$, with J' having no solution above $x = 0$ (J and J' are coprime); in particular, there exist polynomials u, v with $ux^m + v = 1$ and v in J'. From this, we get $J = (ux^m + v)J$, and every element in ux^mJ is a multiple of x^m, while every element in vJ is in $\langle f_1, \dots, f_t\rangle$.

Suppose $k \geq m$. As above, we also have polynomials u', v' with $u'x^{k-m} + v' = 1$ and v' in J'. Multiplying by x^m shows that x^m is in the ideal $H = \langle f_1, \dots, f_t, x^k\rangle$, so that $H = J$ (this proves the last claim in the lemma). In this case, the smallest power of x in H is thus x^m.

Suppose $k \leq m$. In this case, we prove that the minimal power of x in $H = \langle f_1, \dots, f_t, x^k\rangle$ is x^k. First, note that in this case, $H = \langle f_1, \dots, f_t, x^m, x^k\rangle = J + \langle x^k\rangle$, and let x^e be the minimum power of x in H; suppose $e < k$, so that $e < m$. It follows that x^e is the normal form of a polynomial of the form fx^k, modulo the Gröbner basis $\mathcal{G}$ of J. However, Lazard's structure theorem [109, Theorem 1] implies that through reduction modulo such a Gröbner basis, no term of x-degree less than k can appear, a contradiction. □

This allows us to design an algorithm GroebnerBasisAtZero that computes the Gröbner basis of J (under the position assumption in the lemma), even though we do not know m in advance: we call HowellGroebnerBasis with inputs the polynomials $(f_1, \dots, f_t, x^k)$, for $k = 2^i$, with $i = 0, 1, \dots,$ until the output does *not* contain x^k. Indeed, the lemma shows that if x^k is in the Gröbner basis of

$H = \langle f_1, \ldots, f_t, x^k \rangle$, we have $k \le m$, while if it is not, we have reached $k > m$, and the output is the Gröbner basis of J.

Altogether, we do $O(\log(m))$ calls to HowellGroebnerBasis, with always $k \le 2m$. With d the maximum degree of $f_1, \ldots, f_t$, the runtime is $O\tilde{}(tD^\omega m)$ operations in $\mathbb{K}$, with D in $\{\Delta_1(d_y), \Delta_2(d), \Delta_3(d)\}$, depending on our assumptions on $f_1, \ldots, f_t$ (recall that d_y and d are the maximum y-degree, resp. degree, of the input).

5.3 Coefficient size and bad reductions

In this section, we suppose that our base field $\mathbb{K}$ is endowed with a valuation $|\ |_{\mathbb{K}} : \mathbb{K} \to \mathbb{R}_{\ge 0}$, that is, a mapping that satisfies the following properties:

(1). $|\ |_{\mathbb{K}}$ vanishes at zero, and only at zero,

(2). $|uv|_{\mathbb{K}} = |u|_{\mathbb{K}}\ |v|_{\mathbb{K}}$, for u, v in $\mathbb{K}$,

(3). $|u + v|_{\mathbb{K}} \le |u|_{\mathbb{K}} + |v|_{\mathbb{K}}$, for u, v in $\mathbb{K}$.

If the stronger condition $|u + v|_{\mathbb{K}} \le \max(|u|_{\mathbb{K}}, |v|_{\mathbb{K}})$ holds instead of (3), we say that $|\ |_{\mathbb{K}}$ is *ultrametric* otherwise we say $|\ |_{\mathbb{K}}$ is *Archimedean*. We assume that $\mathbb{K}$ is the fraction field of a domain $\mathbb{A}$, and that all nonzero elements in $\mathbb{A}$ have absolute value at least 1. The main examples we have in mind are $\mathbb{K} = \mathbb{Q}$, with $\mathbb{A} = \mathbb{Z}$ and $|\ |_{\mathbb{Q}}$ the usual absolute value, and $\mathbb{K} = k(\boldsymbol{t})$, for k a field, with $\mathbb{A} = k[\boldsymbol{t}]$ and $|f/g|_{k(\boldsymbol{t})} = e^{\deg(f)-\deg(g)}$, for $f, g \ne 0$ (the latter is ultrametric).

The *height* of a nonzero element $u \in \mathbb{K}$ is $\max(0, \log(|u|_{\mathbb{K}}))$; in particular, for $\mathbb{K} = \mathbb{Q}$, the height of $u \in \mathbb{Z} - \{0\}$ is thus $\log(|u|)$, and for $\mathbb{K} = k(\boldsymbol{t})$, the height of a nonzero polynomial $u \in k[\boldsymbol{t}]$ is its degree.

Our goal here is to give height bounds on the elements in the lexicographic Gröbner basis of some polynomials $\mathcal{F} = (f_1, \ldots, f_t)$. The key quantity $H(\mathcal{F})$, together with an element $\beta_{\mathcal{F}} \in \mathbb{A}$, are defined as follows.

❧ **Definition 29.** *Consider polynomials $\mathcal{F} = (f_1, \ldots, f_t)$ in $\mathbb{A}[x, y]$, let I be the ideal they generate in $\mathbb{K}[x, y]$, with lexicographic Gröbner basis $\mathcal{G} = (g_0, \ldots, g_s)$. We define $H(\mathcal{F})$ as the smallest integer such that there exists $\beta_{\mathcal{F}}$ nonzero in $\mathbb{A}$ for which we have:*

- *the polynomials $\beta_{\mathcal{F}} g_0, \ldots, \beta_{\mathcal{F}} g_s$ are in $\mathbb{A}[x, y]$*
- *all coefficients of $\beta_{\mathcal{F}} g_0, \ldots, \beta_{\mathcal{F}} g_s$ (which include in particular $\beta_{\mathcal{F}}$) have height at most $H(\mathcal{F})$*

- *for any maximal ideal* $\mathfrak{m} \subset \mathbb{A}$*, with residual field* $\mathbb{F} = \mathbb{A}/\mathfrak{m}$*, if* $\beta_{\mathcal{F}} \notin \mathfrak{m}$*,* $\mathcal{G} \bmod \mathfrak{m}$ *is the lexicographic Gröbner basis of* $\langle f_1 \bmod \mathfrak{m}, \dots, f_t \bmod \mathfrak{m}\rangle$ *in* $\mathbb{F}[x, y]$.

In order to give upper bounds on $H(\mathcal{F})$, we introduce two functions $B(n, d, h)$ and $C(t, d, D, h)$. The first one, $B(n, d, h)$, is defined by

- $B(n, d, h) = (N + 1)h$ if $|\ |_{\mathbb{K}}$ is ultrametric, with $N = n^2d - nd + n$
- $B(n, d, h) = (N + 1)h + N\log(N) + \log(n(d + 1))$ in general.

Next, $C(t, d, D, h)$ is the function defined by

- $C(t, d, D, h) = B(tD, d, h) + h$ if $|\ |_{\mathbb{K}}$ is ultrametric
- $C(t, d, D, h) = B(tD, d, h) + h + \log(2)$ in general.

Whether $|\ |_{\mathbb{K}}$ is Archimedean or not, $B(n, d, h)$ is in $O\tilde{~}(n^2dh)$ and $C(t, d, D, h)$ is in $O\tilde{~}(t^2D^2dh)$.

Proposition 5.3.1. *Let* $\mathcal{F} = (f_1, \dots, f_t)$ *be in* $\mathbb{A}[x, y]$*, for* $t \geq 2$*, such that the ideal* $I = \langle f_1, \dots, f_t\rangle \subset \mathbb{K}[x, y]$ *has dimension zero. Suppose that all* f_i*'s have degree at most* d *and coefficients of height at most* h*.*

(i) *if there exists* i *in* $\{1, \dots, t\}$ *such that the coefficient of* y^d *in* f_i *is a nonzero constant,* $H(\mathcal{F}) \leq C(t, d, \Delta_1(d), h) \in O\tilde{~}(t^2d^3h)$

(ii) *if* $t = 2$*,* $H(\mathcal{F}) \leq C(2, d, \Delta_2(d), h) \in O\tilde{~}(d^5h)$

(iii) *in general,* $H(\mathcal{F}) \leq C(t, d, \Delta_3(d), h) \in O\tilde{~}(t^2d^9h)$

The proposition will follow from height bounds for Hermite forms of matrices due to Storjohann, which we recall in the first subsection; from this, the extension to lexicographic Gröbner bases follows directly from the discussion in the previous section.

To our knowledge, no previous bounds were given in this setting; however, some results are available for particular cases. We discuss them here in the particular case $\mathbb{K} = \mathbb{Q}$; the results quoted below also cover more general cases.

Several previous results covered the case of *radical* ideals with generators in $\mathbb{Z}[x, y]$ and finitely many solutions. If their Gröbner basis $\mathcal{G}$ is a *triangular set* (that is, $\mathcal{G} = (g_0, g_1)$, with leading terms of the form y^{n_0} and x^{m_s}, respectively), the results in [50] show that the polynomials in $\mathcal{G}$ have coefficients with numerator and denominator of height $O\tilde{~}(d^3h + d^4)$. Our result does not feature the term d^4, but this might be due to the proof techniques of [50], which are not limited to systems in two variables. If we keep the radicality assumption, but allow arbitrary leading terms, the best previous bound we are aware of is $O\tilde{~}(d^7h + d^8)$, from [46].

5.3.1 Coefficient size and bad reductions for Hermite normal forms

We recall here results of Storjohann's [149] on size bounds and unlucky reductions for Hermite normal forms of matrices with entries in $\mathbb{A}[x] \subset \mathbb{K}[x]$. That reference deals with $\mathbb{A} = \mathbb{Z}$, but the same treatment applies to our more general context. We briefly recall the key elements of the proof in [149], skipping the details that can be found in that reference.

Proposition 5.3.2 ([149, Section 6.2])**.** *Let $\boldsymbol{A}$ be in $\mathbb{A}[x]^{n\times n}$, with nonzero determinant and entries of degree at most $d > 0$ and height at most h. Let further $\boldsymbol{H}$ be the Hermite normal form of $\boldsymbol{A}$. Then, there exists α nonzero in $\mathbb{A}$ such that:*

- *all entries of $\alpha\boldsymbol{H}$ are in $\mathbb{A}[x]$*
- *α and the coefficients of all entries of $\alpha\boldsymbol{H}$ have height at most $B(n, d, h)$*
- *for any maximal ideal $\mathfrak{m} \subset \mathbb{A}$, with residual field $\mathbb{F} = \mathbb{A}/\mathfrak{m}$, if $\alpha \notin \mathfrak{m}$, then $\boldsymbol{H} \bmod \mathfrak{m}$ is the Hermite normal form of $\boldsymbol{A} \bmod \mathfrak{m}$ in $\mathbb{F}[x]^{n\times n}$.*

Sketch of proof. Since $\boldsymbol{A}$ is invertible over $\mathbb{K}(x)$, the transformation matrix $\boldsymbol{U}$ such that $\boldsymbol{H} = \boldsymbol{A}\boldsymbol{U}$ is uniquely defined, and it has entries of degree at most $D = (n-1)d$.

Storjohann showed how to linearize the computation of $\boldsymbol{U}$. Set $N = n(D+1) \le n^2 d$; then, there exist $N' \le N$ and matrices $\mathcal{G}_{\text{lin}}, \boldsymbol{A}_{\text{lin}}, \boldsymbol{U}_{\text{lin}}$ with entries in $\mathbb{K}$ and of respective sizes $N' \times n$, $N' \times N'$ and $N' \times n$ such that

- $\mathcal{G}_{\text{lin}} = \boldsymbol{A}_{\text{lin}}\boldsymbol{U}_{\text{lin}}$,
- $\mathcal{G}_{\text{lin}}$ has exactly one nonzero entry per column, which is 1,
- $\boldsymbol{A}_{\text{lin}}$ is invertible, and its entries are coefficients of the entries of $\boldsymbol{A}$,
- for $1 \le i \le n$, the entries on the ith row of $\boldsymbol{U}_{\text{lin}}$ are the coefficients of degrees $0, \dots, D$ of $U_{i,1}$, then of $U_{i,2}$, $\dots$, and finally of $U_{i,n}$.

Let $\alpha \in \mathbb{A} - \{0\}$ be the determinant of $\boldsymbol{A}_{\text{lin}}$. The previous items show that $\alpha\boldsymbol{U}$ is in $\mathbb{A}[x]^{n\times n}$, and the relation $\boldsymbol{H} = \boldsymbol{A}\boldsymbol{U}$ shows that is also the case for $\alpha\boldsymbol{H}$.

Let $\mathfrak{m}$ be a maximal ideal in $\mathbb{A}$ such that $\alpha \notin \mathfrak{m}$, with residual field $\mathbb{F} = \mathbb{A}/\mathfrak{m}$. We deduce from the above that $\boldsymbol{H}$ and $\boldsymbol{U}$ are in $\mathbb{A}_{\mathfrak{m}}[x]^{n\times n}$. If we let $\bar{\boldsymbol{H}}$, $\bar{\boldsymbol{A}}$ and $\bar{\boldsymbol{U}}$ be the reductions of all these matrices modulo $\mathfrak{m}$, we see that we have $\bar{\boldsymbol{H}} = \bar{\boldsymbol{A}}\bar{\boldsymbol{U}}$ in $\mathbb{F}[x]^{n\times n}$, and since $\bar{\boldsymbol{H}}$ is still in Hermite normal form, and $\bar{\boldsymbol{U}}$ still invertible, $\bar{\boldsymbol{H}}$ is the Hermite form of $\bar{\boldsymbol{A}}$. It remains to give a bound on the height of α and of the coefficients of the entries of $\alpha\boldsymbol{H}$.

- The coefficient α is the determinant of the matrix $\boldsymbol{A}_{\mathrm{lin}}$ of size at most N, with entries of height at most h; the entries of $\alpha \boldsymbol{U}_{\mathrm{lin}}$ are minors of $\boldsymbol{A}_{\mathrm{lin}}$. If $|\ |_{\mathbb{K}}$ is ultrametric, their height is thus at most Nh; in general, the bound is $Nh + N\log(N)$.

- The matrix $\alpha\boldsymbol{H}$ is the product of $\alpha\boldsymbol{U}$ and $\boldsymbol{A}$. If $|\ |_{\mathbb{K}}$ is ultrametric, the former has polynomial entries with coefficients of height at most Nh, whereas the bound is h for the latter, so the entries of $\alpha\boldsymbol{H}$ have coefficients of height at most $(N+1)h$. For a general $|\ |_{\mathbb{K}}$, we have to take into account the degree of $\alpha\boldsymbol{U}$ and $\boldsymbol{A}$, respectively at most $D = (n-1)d$ and d. As a result, the height bound on the coefficients of the entries of $\alpha\boldsymbol{H}$ is $(N+1)h+N\log(N)+\log(n(d+1))$. □

5.3.2 Application to Gröbner bases

Let $f_1, \dots, f_t$ be as in Proposition 5.3.1. First, we define an element $\gamma \in \mathbb{A}$ and integer D through the following case discussion:

- If we are in case (i), we know that at least one of the f_i's has a coefficient of degree d (in y) in $\mathbb{A} - \{0\}$; let γ be such a coefficient. We let $D = \Delta_1(d)$ from Proposition 5.2.2.

- in case (ii) or (iii), we let $\gamma = 1$, and we take respectively $D = \Delta_2(d)$ or $D = \Delta_3(d)$, with notation from Proposition 5.2.2.

In any case, we know that $\Delta(\mathcal{F}) \le D$, so we can apply Proposition 5.2.1; it shows that we can recover the (minimal, reduced) lexicographic Gröbner basis of $I = \langle f_1, \dots, f_t\rangle$ from the columns of the Hermite form of the Sylvester-like matrix $\boldsymbol{S}$ defined in that proposition.

As in the previous section, there is a $(d+D)\times tD$ matrix $\boldsymbol{S}'$ obtained by permuting the columns of $\boldsymbol{S}$ whose leading $(d+D)\times(d+D)$ minor is nonzero. Consider again the $tD \times tD$ square matrix $\boldsymbol{S}^{\mathrm{sq}}$ of Eq. (5.2.0.1) and its Hermite form $\boldsymbol{H}^{\mathrm{sq}}$; the first $d+D$ rows of $\boldsymbol{H}^{\mathrm{sq}}$ are the Hermite form $\boldsymbol{H}$ of $\boldsymbol{S}$.

Since $\boldsymbol{S}^{\mathrm{sq}}$ has nonzero determinant, we can let α be the non-zero element in $\mathbb{A}$ associated to it by means of Proposition 5.3.2, and we let $\beta = \alpha\gamma$. That proposition shows that all entries of $\beta\boldsymbol{H}^{\mathrm{sq}}$, and thus of $\beta\boldsymbol{H}$, have entries in $\mathbb{A}[x]$, the latter having coefficients of height at most $C(t, d, D, h)$ (this includes in particular β). By means of Proposition 5.2.1, we deduce that these height bounds apply in particular to the Gröbner basis $\boldsymbol{G} = (g_0, \dots, g_s)$ of I.

Suppose then that $\mathfrak{m} \subset \mathbb{A}$ is a maximal ideal that does not contain β. Then, because α is not in $\mathfrak{m}$, Proposition 5.3.2 shows that $\bar{\boldsymbol{H}}^{\mathrm{sq}} = \boldsymbol{H}^{\mathrm{sq}} \bmod \mathfrak{m}$ is the Hermite normal form of $\bar{\boldsymbol{S}}^{\text{sq}} = \boldsymbol{S}^{\text{sq}} \bmod \mathfrak{m}$. Considering only the first tD rows, we see that $\bar{\boldsymbol{H}} = \boldsymbol{H} \bmod \mathfrak{m}$ is the Hermite normal form of $\bar{\boldsymbol{S}} = \boldsymbol{S} \bmod \mathfrak{m}$. Now, let us prove that we still have $\Delta(\bar{\mathcal{F}}) \leq D$.

- If we are in case (i), since γ is not in $\mathfrak{m}$, at least one of the polynomials $\bar{f}_i = f_i \bmod \mathfrak{m}$ has its coefficient of y-degree d a nonzero constant in $\mathbb{F} = \mathbb{A}/\mathfrak{m}$. Since all $\bar{f}_i$'s have degree at most d, we deduce $\Delta(\bar{\mathcal{F}}) = d$ in this case (first item of Proposition 5.2.2)
- If we are in case (ii) or (iii), the discussion above shows that $\bar{g}_0$ and $\bar{g}_s$ are in the ideal $\langle \bar{f}_1, \dots, \bar{f}_t \rangle$, so that this ideal admits finitely many solutions in an algebraic closure of the residual field $\mathbb{F} = \mathbb{A}/\mathfrak{m}$. Using the second and third items of Proposition 5.2.2 gives our claim.

We can then apply Proposition 5.2.1 to $\bar{\mathcal{F}} = (\bar{f}_1, \dots, \bar{f}_t)$, and deduce that the columns of the Hermite form of $\bar{\boldsymbol{S}}$ give a detaching basis, and in particular the lexicographic Gröbner basis of $\langle \bar{f}_1, \dots, \bar{f}_t \rangle$. This proves the proposition.

5.4 Applying generic changes of coordinates

In this section, we work over a base field $\mathbb{K}$, and we quantify changes of coordinates that ensure three desirable properties: curves in Noether position, one-to-one projections and Borel-fixed-ness of the initial ideal. For our discussion here, it will be convenient to consider changes of coordinates with entries in $\overline{\mathbb{K}}$ (and thus to work in $\overline{\mathbb{K}}[x, y]$), but the algorithms will take them with entries in $\mathbb{K}$.

We write $\boldsymbol{\gamma}$ for a 2×2 matrix $\boldsymbol{\gamma} = [\gamma_{i,j}]_{1 \leq i,j \leq 2}$ with entries in $\overline{\mathbb{K}}$, and we identify $M_2(\overline{\mathbb{K}})$ with $\overline{\mathbb{K}}^4$ through $\boldsymbol{\gamma} \mapsto [\gamma_{1,1}, \gamma_{1,2}, \gamma_{2,1}, \gamma_{2,2}]$. For γ in $\mathrm{GL}_2(\overline{\mathbb{K}})$ as above and f in $\overline{\mathbb{K}}[x, y]$, we write $f^{\boldsymbol{\gamma}} = f(\gamma_{1,1}x + \gamma_{2,1}y, \gamma_{1,2}x + \gamma_{2,2}y)$. Note that for two matrices $\boldsymbol{\gamma}, \boldsymbol{\gamma}'$, we have $(f^{\boldsymbol{\gamma}})^{\boldsymbol{\gamma}'} = f^{\boldsymbol{\gamma}'\boldsymbol{\gamma}}$, so $\mathrm{GL}_2(\overline{\mathbb{K}})$ acts on the left on $\overline{\mathbb{K}}[x, y]$.

5.4.1 Equations in general position

For $\mathcal{F} = (f_1, \dots, f_t)$ as in the previous sections, the best degree and height bounds $\Delta(\mathcal{F})$ and $H(\mathcal{F})$ apply when the input equations have a particular property: at least one f_i has a term of maximal degree that involves y only. Geometrically, this means that the curve $V(f_i) \subset \overline{\mathbb{K}}^2$ has no vertical asymptote; we also say that it is in *Noether position.* The following lemma is straightforward.

Proposition 5.4.1. *Take f in $\mathbb{K}[x,y]$ of degree d. Then there exists a hypersurface $X \subset \overline{\mathbb{K}}^4$ of degree at most d such that if $\boldsymbol{\gamma}$ is in $\overline{\mathbb{K}}^4 - X$, the coefficient of y^d in $f^{\boldsymbol{\gamma}}$ is nonzero.*

Proof. Let $f_d \in \mathbb{K}[x,y]$ be the homogeneous component of degree d in f. Then the coefficient of y^d in $f^{\boldsymbol{\gamma}}$ is $f_d(\gamma_{2,1}, \gamma_{2,2})$. □

Another favourable situation, illustrated when we dealt with Howell forms, occurs when the projection $V(\langle \mathcal{F} \rangle) \to \overline{\mathbb{K}}$ given by $(\alpha, \beta) \mapsto \alpha$ is one-to-one. Again, the proof is standard.

Proposition 5.4.2. *Let $\mathcal{F} = (f_1, \dots, f_t)$ be in $\mathbb{K}[x,y]$ of degrees at most d, and suppose that $V(\langle \mathcal{F} \rangle)$ is finite. Then there exists a hypersurface $X \subset \overline{\mathbb{K}}^4$ of degree at most d^4 such that if $\boldsymbol{\gamma}$ is invertible and in $\overline{\mathbb{K}}^4 - X$, the projection on the first factor $V(\langle \mathcal{F}^{\boldsymbol{\gamma}} \rangle) \to \overline{\mathbb{K}}$ is one-to-one.*

Proof. Since we assume that the zero-set $V(\mathcal{F})$ is finite, its cardinal D is at most d^2, by [89, Proposition 2.3]; we write $V(\mathcal{F}) = (\alpha_i, \beta_i)_{1 \le i \le D}$.

For $\boldsymbol{\gamma}$ invertible of determinant $g \neq 0$, the zero-set $V(\mathcal{F}^{\boldsymbol{\gamma}})$ are the point of coordinates $((\gamma_{2,2}\alpha_i - \gamma_{2,1}\beta_i)/g, (-\gamma_{1,2}\alpha_i + \gamma_{1,1}\beta_i)/g)$. It follows that the projection $V(\mathcal{F}^{\boldsymbol{\gamma}}) \to \overline{\mathbb{K}}$ is one-to-one if and only if, for $1 \le i < j \le D$, we have $\gamma_{2,2}(\alpha_i - \alpha_j) - \gamma_{2,1}(\beta_i - \beta_j) \neq 0$. Since the vector $(\alpha_i - \alpha_j, \beta_i - \beta_j)$ is nonzero, this imposes a linear constraint on $\boldsymbol{\gamma}$. There are $D(D-1)/2 \le D^2$ pairs i, j to consider, and the conclusion follows. □

5.4.2 The initial ideal is Borel-fixed

The second property we consider concerns the initial ideal $\mathbf{In}(I)$ of an ideal $I \subset \overline{\mathbb{K}}[x,y]$, respective to a monomial order $\prec$ for which $x \prec y$. We say that an ideal $J \subset \overline{\mathbb{K}}[x,y]$ is *Borel-fixed* if it is stable under the action of the group of lower-diagonal invertible matrices (this differs from the convention in e.g. [61, Chapter 15], which uses upper-triangular matrices; this is because we choose $x \prec y$ rather than $y \prec x$).

Galligo proved that for homogeneous ideals in multivariate power series rings (endowed with local orders), initial ideals are Borel-fixed in generic coordinates [72]. Similar statements hold in polynomial rings, but all published proofs we are aware of apply to homogeneous ideals, and none of them gives a quantitative statement on the "degree of genericity". In this subsection, we prove such a statement for I of dimension zero, but without the homogeneity assumption. We adapt the proof for the homogeneous case given in [61], using the dimension zero assumption to dispense with the use of Dickson's lemma. While the proof is given in the bivariate context of this paper, it applies without modification in more than two variables.

For $S \subset \overline{\mathbb{K}}[x,y]$ and $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$, we let $S^{\boldsymbol{\gamma}} = \{f^{\boldsymbol{\gamma}} \mid f \in S\}$. If S is a $\overline{\mathbb{K}}$-vector space, resp. an ideal, $S^{\boldsymbol{\gamma}}$ is a $\overline{\mathbb{K}}$-vector space of the same dimension as S (resp. an ideal).

Proposition 5.4.3. *Let $I \subset \overline{\mathbb{K}}[x,y]$ be an ideal of dimension zero, and let $\delta = \dim_{\overline{\mathbb{K}}}(\overline{\mathbb{K}}[x,y]/I)$. Then, there exists a hypersurface $\mathcal{F}_3 \subset \overline{\mathbb{K}}^4$ of degree at most $\delta^3 + 3$ such that if $\boldsymbol{\gamma}$ is in $\overline{\mathbb{K}}^4 - \mathcal{F}_3$, $\boldsymbol{\gamma}$ is invertible and the initial ideal of $I^{\boldsymbol{\gamma}}$ is Borel-fixed.*

Before proving the proposition, we point out the main consequence we will derive from it, regarding the shape the Gröbner basis $\mathcal{G} = (g_0, \dots, g_s)$ of $I^{\boldsymbol{\gamma}}$ (as usual, we list them in decreasing order). For any $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$, the minimal monomial generators of $\mathbf{In}(I^{\boldsymbol{\gamma}})$ all have total degree at most δ. Thus, if $\mathbb{K}$ has characteristic either zero or greater than δ, Theorem 15.23 in [61] shows that if $\mathbf{In}(I^{\boldsymbol{\gamma}})$ is Borel-fixed, g_i has y-degree $s-i$, for $i = 0, \dots, s$. This is a favourable situation from the computational point of view; we will return to these aspects in the next section.

The proof of the proposition occupies the rest of this section. In what follows, the monomial order $\prec$ and the ideal I are fixed; the initial term of a nonzero $f \in \overline{\mathbb{K}}[x,y]$ is written $\mathrm{in}(f)$. We define the following:

- For $d \geq 0$, we write $I_{\leq d} = I \cap \overline{\mathbb{K}}[x,y]_{\leq d}$. One readily checks that for $\boldsymbol{\gamma}$ in $\mathrm{GL}_n(\overline{\mathbb{K}})$, $(I^{\boldsymbol{\gamma}})_{\leq d} = (I_{\leq d})^{\boldsymbol{\gamma}}$, so we simply write this set $I^{\boldsymbol{\gamma}}_{\leq d}$.

- $\mathbf{In}(I)$ is the initial ideal of I for the order $\prec$.

- For any $\overline{\mathbb{K}}$-vector space $S \subset \overline{\mathbb{K}}[x,y]$, we let $\mathbf{in}(S)$ be the $\overline{\mathbb{K}}$-vector space spanned by all $\mathrm{in}(f)$, for f in S.

As [61], we introduce the exterior algebra $\wedge(\overline{\mathbb{K}}[x,y])$ in order to describe the action of $\mathrm{GL}_2(\overline{\mathbb{K}})$ on vector subspaces in $\overline{\mathbb{K}}[x,y]$. A nonzero exterior product $m_1 \wedge \cdots \wedge m_{s_d}$, with all m_i's pairwise distinct monomials, admits a *normal form*, obtained by reordering all m_i's in decreasing order. Two such expressions are compared using the lexicographic order on their normal forms.

Lemma 5.4.1. *Let $S \subset \overline{\mathbb{K}}[x,y]$ be a vector space of finite dimension s. Then* $\mathbf{in}(S)$ *has a uniquely defined monomial basis $(n_1, \dots, n_s)$ with $n_1 > \cdots > n_s$, and for any basis $(g_1, \dots, g_s)$ of S, the maximal term in $g_1 \wedge \cdots \wedge g_s$ is $c n_1 \wedge \cdots \wedge n_s$, for some non-zero constant $c \in \overline{\mathbb{K}}$.*

Proof. Let $(f_1, \ldots, f_s)$ be a $\overline{\mathbb{K}}$-basis of S. Without loss of generality, assume that f_1 has the maximal leading term. By linear combinations, we can further assume that $f_2, \ldots, f_s$ have leading terms less than that of f_1. Continuing this way, we end up with generators $f_1, \ldots, f_s$ of S with leading monomials $n_1 > \cdots > n_s$.

By definition, these monomials are all in $\mathbf{in}(S)$, and they are linearly independent. Conversely, if we take f in $\mathbf{in}(S)$, we have $f = \sum_{i \in B} c_i \mathrm{in}(h_i)$, for some h_i in S. The leading term of any (nonzero) h_i must be one of $n_1, \ldots, n_s$, so f is in the $\overline{\mathbb{K}}$-span of $n_1, \ldots, n_s$. This proves that $\{n_1, \ldots, n_s\}$ is a $\mathbb{K}$-basis of $\mathbf{in}(S)$ (and thus, necessarily its unique monomial basis).

For the second claim, expanding the product shows that the leading term in $f_1 \wedge \cdots \wedge f_s$ is $kn_1 \wedge \cdots \wedge n_s$, for some nonzero $k \in \overline{\mathbb{K}}$. Now, for any other basis $(g_1, \ldots, g_s)$, $f_1 \wedge \cdots \wedge f_s = \alpha g_1 \wedge \cdots \wedge g_s$, for some nonzero $\alpha \in \overline{\mathbb{K}}$; the conclusion follows. □

We call the monomial basis $(n_1, \ldots, n_s)$ in this lemma, sorted in decreasing order, the *canonical basis* of $\mathbf{in}(S)$.

Let further $\mathbf{\Gamma} = [a_{i,j}]_{1 \le i,j \le 2}$ be a 2×2 matrix with indeterminate entries. For $d \geq 0$, let $s_d = \dim_{\overline{\mathbb{K}}}(I_{\le d})$, take a $\overline{\mathbb{K}}$-basis $f_{d,1}, \ldots, f_{d,s_d}$ of $I_{\le d}$, and consider $f_{d,1}^{\mathbf{\Gamma}}, \ldots, f_{d,s_d}^{\mathbf{\Gamma}}$ in $\overline{\mathbb{K}}[\boldsymbol{a}][x, y]$.

Lemma 5.4.2. *The maximal term in $f_{d,1}^{\mathbf{\Gamma}} \wedge \cdots \wedge f_{d,s_d}^{\mathbf{\Gamma}}$ has the form $C_d n_{d,1} \wedge \cdots \wedge n_{d,s_d}$, for C_d a nonzero polynomial of degree at most ds_d in $\overline{\mathbb{K}}[\boldsymbol{a}]$ and monomials $n_{d,1} > \cdots > n_{d,s_d}$.*

Proof. Replacing $\mathbf{\Gamma}$ by the 2×2 identity matrix gives $f_{d,1} \wedge \cdots \wedge f_{d,s_d}$, which is nonzero, so $f_{d,1}^{\mathbf{\Gamma}} \wedge \cdots \wedge f_{d,s_d}^{\mathbf{\Gamma}}$ itself is nonzero, and thus has a leading term of the claimed form. Each $f_{d,i}$ has degree at most d in x, y, so $f_{d,i}^{\mathbf{\Gamma}}$ has degree at most d in $\boldsymbol{a}$ and the degree bound on C_d follows. □

Lemma 5.4.3. *The following holds:*

- *For any $\boldsymbol{\gamma}$ in $M_2(\overline{\mathbb{K}})$ and any $g_1, \ldots, g_{s_d}$ in $I_{\le d}^{\boldsymbol{\gamma}}$, all monomials in $g_1 \wedge \cdots \wedge g_{s_d}$ are less than or equal to $n_{d,1} \wedge \cdots \wedge n_{d,s_d}$.*
- *If $\boldsymbol{\gamma} \in \mathrm{GL}_2(\overline{\mathbb{K}})$ does not cancel C_d, $(n_{d,1}, \ldots, n_{d,s_d})$ is the canonical $\overline{\mathbb{K}}$-basis of $\mathbf{in}(I_{\le d}^{\boldsymbol{\gamma}})$.*

Proof. First item: assume $g_1, \ldots, g_{s_d}$ are linearly independent (otherwise, the wedge product is zero). Then, they form a $\overline{\mathbb{K}}$-basis of $I_{\le d}^{\boldsymbol{\gamma}}$, and it follows that $g_1 \wedge \cdots \wedge g_{s_d} = k f_{d,1}^{\boldsymbol{\gamma}} \wedge \cdots \wedge f_{d,s_d}^{\boldsymbol{\gamma}}$, for some non-zero constant k in $\overline{\mathbb{K}}$. So the terms in $g_1 \wedge \cdots \wedge g_{s_d}$ are obtained by evaluating those of $f^{\mathbf{\Gamma}}_{d,1} \wedge \cdots \wedge f^{\mathbf{\Gamma}}_{d,s_d}$ at the entries of $\boldsymbol{\gamma}$, and the conclusion follows from the definition of $n_{d,1}, \dots, n_{d,s_d}$.

Second item: the assumption implies that the maximal term in $f^{\boldsymbol{\gamma}}_{d,1} \wedge \cdots \wedge f^{\boldsymbol{\gamma}}_{d,s_d}$ is $cn_{d,1} \wedge \cdots \wedge n_{d,s_d}$, for c non-zero in $\overline{\mathbb{K}}$. Since $f^{\boldsymbol{\gamma}}_{d,1}, \dots, f^{\boldsymbol{\gamma}}_{d,s_d}$ are a $\overline{\mathbb{K}}$-basis of $I^{\boldsymbol{\gamma}}_{\le d}$, Lemma 5.4.1 shows that $(n_{d,1}, \dots, n_{d,s_d})$ is the canonical basis of $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d})$. □

For $d \ge 0$, let $B_d \subset \overline{\mathbb{K}}[x, y]$ be the $\overline{\mathbb{K}}$-span of $n_{d,1}, \dots, n_{d,s_d}$. By the previous lemma, if $C_d(\boldsymbol{\gamma}) \ne 0$, $B_d = \mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d})$.

Lemma 5.4.4. *For $d \ge 0$, $B_d \subset B_{d+1}$.*

Proof. We first prove that each $n_{d,i}$ is in B_{d+1}. Take $\boldsymbol{\gamma} \in \mathrm{GL}_2(\overline{\mathbb{K}})$ that cancels neither C_d nor C_{d+1}. Then, we saw that $n_{d,i}$ is in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d})$, so it is a linear combination $\sum_j \mathrm{in}(f_j)$, for some f_j in $I^{\boldsymbol{\gamma}}_{\le d}$, and so, in fact, $n_{d,i} = \mathrm{in}(f)$ for some f in $I^{\boldsymbol{\gamma}}_{\le d}$. Then, f is in $I^{\boldsymbol{\gamma}}_{\le d+1}$, so $n_{d,i}$ is in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d+1})$. By assumption on $\boldsymbol{\gamma}$, $n_{d,i}$ is thus in B_{d+1}. Because B_d and B_{d+1} are $\overline{\mathbb{K}}$-vector spaces, this proves $B_d \subset B_{d+1}$. □

Let $B = \cup_{d \ge 0} B_d$. Note that by the previous lemma, for any $D \ge 0$, we have $B = \cup_{d \ge D} B_d$.

Lemma 5.4.5. *B is a monomial ideal.*

Proof. Let f, g be in B. Then (Lemma 5.4.4) there is d such that f and g are in B_d. B_d is a vector space, so for any u, v in $\overline{\mathbb{K}}$, $uf + vg$ is in B_d, and thus in B. So B is a $\overline{\mathbb{K}}$-vector space.

Next, we prove that $x_j B_d$ is contained in B_{d+1}, for $d \ge 0$ and j in $\{1, \dots, n\}$. Take $\boldsymbol{\gamma}$ that cancels neither C_d nor C_{d+1}. As in the previous lemma, $n_{d,i}$ is of the form $n_{d,i} = \mathrm{in}(f)$ for some f in $I^{\boldsymbol{\gamma}}_{\le d}$. Now, $x_j f$ is in $I^{\boldsymbol{\gamma}}_{\le d+1}$, so its initial term $x_j n_{d,i}$ is in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d+1})$. By assumption on $\boldsymbol{\gamma}$, $x_j n_{d,i}$ is thus in B_{d+1}. By additivity, $x_j B_d$ is contained in B_{d+1}.

As a result, for any monomial m of degree e, mB_d is contained in B_{d+e} (by induction), and thus in B. It follows that mB is contained in B, so B is an ideal.

Finally, let $M \subset B$ be the union of all sets $\{n_{d,1}, \dots, n_{d,s_d}\}$, for $d \ge 0$. Let f be in B, so that f is in B_d for some $d \ge 0$. Since B_d is generated by $\{n_{d,1}, \dots, n_{d,s_d}\}$ as a vector space, f is in the $\overline{\mathbb{K}}$-span of M. Thus, M generates B as a vector space, and then also as an ideal, so that B is a monomial ideal. □

The next lemmas prove that for generic $\boldsymbol{\gamma}$, B is the initial ideal of $I^{\boldsymbol{\gamma}}$.

Lemma 5.4.6. *For $d \ge 0$ and $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$, $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d}) \subset \mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d}$.*

Proof. Take $f = \sum_i \mathrm{in}(f_i)$ in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d})$, with all f_i's in $I^{\boldsymbol{\gamma}}_{\le d}$. Then, all f_i's are in $I^{\boldsymbol{\gamma}}$, so f is in $\mathbf{In}(I^{\boldsymbol{\gamma}})$. On the other hand, all f_i's, and thus all $\mathrm{in}(f_i)$'s, have degree at most d, so f is in $\mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d}$. □

Lemma 5.4.7. *The ideal B has degree at least $\delta = \deg(I)$.*

Proof. Let $h_1, \ldots, h_t$ be ideal generators of B. Since each h_i belongs to some B_{d_i}, and the sequence $(B_d)_{d \ge 0}$ is increasing (Lemma 5.4.4), there exists $D \ge 0$ such that all h_i's are in B_D.

Take $\boldsymbol{\gamma}$ that does not cancel C_D; then, $B_D = \mathbf{in}(I^{\boldsymbol{\gamma}}_{\le D})$, so that all h_i's are in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le D})$. By Lemma 5.4.6, they are in $\mathbf{In}(I^{\boldsymbol{\gamma}})_{\le D}$, and thus in $\mathbf{In}(I^{\boldsymbol{\gamma}})$. As a result, the whole ideal B is in $\mathbf{In}(I^{\boldsymbol{\gamma}})$, which implies $\deg(B) \ge \deg(I^{\boldsymbol{\gamma}}) = \deg(I)$. □

Lemma 5.4.8. *For $d \ge \delta$ and $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$, $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d}) = \mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d}$.*

Proof. We proved in Lemma 5.4.6 that we have the inclusion $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d}) \subset \mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d}$, for $d \ge 0$ and any $\boldsymbol{\gamma}$. Now, we prove that for $d \ge \delta$ and any $\boldsymbol{\gamma}$, $\dim_{\overline{\mathbb{K}}}(\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le d})) = \dim_{\overline{\mathbb{K}}}(\mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d})$. The former dimension is equal to $\dim_{\overline{\mathbb{K}}}(I^{\boldsymbol{\gamma}}_{\le d})$, by Lemma 5.4.1. Now, for any $\boldsymbol{\gamma}$, both $I^{\boldsymbol{\gamma}}$ and $\mathbf{In}(I^{\boldsymbol{\gamma}})$ have dimension zero and degree δ, so for $d \ge \delta$, $\dim_{\overline{\mathbb{K}}}(I^{\boldsymbol{\gamma}}_{\le d}) = \dim_{\overline{\mathbb{K}}}(\mathbf{In}(I^{\boldsymbol{\gamma}})_{\le d}) = (\delta+1)(\delta+2)/2 - \delta$. □

Lemma 5.4.9. *For $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$ that does not cancel C_δ, $\mathbf{In}(I^{\boldsymbol{\gamma}}) = B$.*

Proof. Take any $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(\overline{\mathbb{K}})$. The ideal $I^{\boldsymbol{\gamma}}$ has degree δ, and thus so does $\mathbf{In}(I^{\boldsymbol{\gamma}})$. The minimal monomial generating set of $\mathbf{In}(I^{\boldsymbol{\gamma}})$, say $g_1, \ldots, g_m$, is thus made of monomials of degree at most δ. So each g_i is in $\mathbf{In}(I^{\boldsymbol{\gamma}})_{\le \delta}$, and thus in $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le \delta})$, by Lemma 5.4.8.

If we suppose that $\boldsymbol{\gamma}$ does not cancel C_δ, then $\mathbf{in}(I^{\boldsymbol{\gamma}}_{\le \delta}) = B_\delta$, so that each g_i is in B_δ, and thus in B. This proves the inclusion $\mathbf{In}(I^{\boldsymbol{\gamma}}) \subset B$, and in particular $\deg(B) \le \deg(\mathbf{In}(I^{\boldsymbol{\gamma}})) = \delta$. Since we saw that $\deg(B) \ge \delta$ (Lemma 5.4.7), these two monomial ideals have the same degree δ, and thus they are equal. □

To prove Proposition 5.4.3, we define $\mathcal{F}_3$ as the vanishing set of either C_δ or the determinant $\gamma_{1,1}\gamma_{2,2} - \gamma_{2,1}\gamma_{1,2}$. We know that C_δ has degree at most δs_δ, with s_δ the dimension of $I_{\le \delta}$. This gives $s_\delta = (\delta+1)(\delta+2)/2 - \delta$, and the degree bound $\deg(\mathcal{F}_3) \le \delta^3 + 3$.

Finally, we establish that B is Borel-fixed; this part of the proof is very close to that of [61, Theorem 15.20].

Lemma 5.4.10. *B is Borel-fixed.*

Proof. We prove that for any matrix $\boldsymbol{I}+\boldsymbol{\eta}$, with $\boldsymbol{\eta}$ having only one entry, that lies under the diagonal, we have $B^{\boldsymbol{I}+\boldsymbol{\eta}} = B$. It is enough to prove that $(B_d)^{\boldsymbol{I}+\boldsymbol{\eta}} = B_d$ for $d \geq 0$ (taking the union will give the conclusion).

Take $d \geq 0$ and recall that $(n_{d,1}, \dots, n_{d,s_d})$ is the (unique, up to permutation) monomial basis of B_d. The polynomials $(n_{d,1}^{\boldsymbol{I}+\boldsymbol{\eta}}, \dots, n_{d,s_d}^{\boldsymbol{I}+\boldsymbol{\eta}})$ are then a basis of $B_d^{\boldsymbol{I}+\boldsymbol{\eta}}$; we will prove that $n_{d,1}^{\boldsymbol{I}+\boldsymbol{\eta}} \wedge \cdots \wedge n_{d,s_d}^{\boldsymbol{I}+\boldsymbol{\eta}} = n_{d,1} \wedge \cdots \wedge n_{d,s_d}$; this implies our claim that $(B_d)^{\boldsymbol{I}+\boldsymbol{\eta}} = B_d$. Write $n = n_{d,1} \wedge \cdots \wedge n_{d,s_d}$, and suppose that $n^{\boldsymbol{I}+\boldsymbol{\eta}}$ is different from n. Then, because $\boldsymbol{\eta}$ is strictly lower triangular, all new terms are greater than n (we are replacing x by $x + gy$, for some constant g). We want to prove that there are no such new terms, so we let $n' > n$ be one of them and derive a contradiction.

Let $\boldsymbol{\gamma}$ be in $\mathrm{GL}_2(\overline{\mathbb{K}})$ that does not cancel C_d, so that $B_d = \mathbf{in}(I_{\leq d}^{\boldsymbol{\gamma}})$. Let $g_1, \dots, g_{s_d}$ be a basis of $I_{\leq d}^{\boldsymbol{\gamma}}$; without loss of generality, we can then assume that they have pairwise distinct leading terms $n_{d,1}, \dots, n_{d,s_d}$. If we let $g = g_1 \wedge \cdots \wedge g_{s_d}$, then for a diagonal matrix $\boldsymbol{\phi}$ with diagonal entries ϕ_1, ϕ_2, the coefficient of n' in the expansion of $g^{(\boldsymbol{I}+\boldsymbol{\eta})\boldsymbol{\phi}}$ is a nonzero polynomial A in ϕ_1, ϕ_2 (this calculation is in the end of the proof of [61, Theorem 15.20]).

Choose ϕ_i's in $\overline{\mathbb{K}}$ such that $A(\phi_1, \phi_2)$ is nonzero and let $h_i = g_i^{(\boldsymbol{I}+\boldsymbol{\eta})\boldsymbol{\phi}}$ for $i = 1, \dots, s_d$, so that $h = h_1 \wedge \cdots \wedge h_{s_d}$ is equal to $g^{(\boldsymbol{I}+\boldsymbol{\eta})\boldsymbol{\phi}}$. By construction, h has a term greater than $n = n_{d,1} \wedge \cdots \wedge n_{d,s_d}$ in its expansion. On the other hand, if we write $\boldsymbol{\gamma}' = (\boldsymbol{I}+\boldsymbol{\eta})\boldsymbol{\phi}\boldsymbol{\gamma}$, we obtain that all h_i's are in $I_{\leq d}^{\boldsymbol{\gamma}'}$. This contradicts the first item in Lemma 5.4.3. □

5.5 Main algorithms

We can finally present our main algorithms, where we use Newton iteration to compute lexicographic Gröbner bases: we are given $\mathcal{F} = (f_1, \dots, f_t)$ in $\mathbb{A}[x, y]$, where $\mathbb{A}$ is domain contained in a field $\mathbb{K}$, and we compute either the Gröbner basis $\mathcal{G} = (g_0, \dots, g_s)$ of $I = \langle f_1, \dots, f_t \rangle$, or the Gröbner basis $\mathcal{G}^0 = (g_0^0, \dots, g_r^0)$ of the $\langle x, y \rangle$-primary component of I using $\mathfrak{m}$-adic approximation, for a maximal ideal $\mathfrak{m}$ of $\mathbb{A}$. In what follows, we give details for the computation of $\mathcal{G}$, assuming there exists a fraction reconstruction algorithm and a valuation for $\mathbb{K}$. We first write the development for $\mathbb{A} = \mathbb{Z}$ and $\mathbb{K} = \mathbb{Q}$; we will mention what modifications are needed if we want to work over $\mathbb{A} = k[\boldsymbol{t}]$. We give the general algorithm after those two examples. Then, we show the adaptation to get an algorithm for $\mathcal{G}_0$.

The algorithms are randomized, it takes a parameter $P \geq 1$; our goal being to obtain the correct output with probability at least $1 - 1/2^P$. Throughout, we make the following assumptions:

- f_1 has maximum degree among the f_i's; we write $d = \deg(f_1)$,
- I has dimension zero.

In terms of notation, we let $\delta = \deg(I) = \dim_{\mathbb{K}} \mathbb{K}[x, y]/I$, so that $\delta \leq d^2$. The other important quantity is the size of the output: we let b be the maximum height of the numerators and denominators of the coefficients in $\mathcal{G}$. Each polynomial in $\mathcal{G}$ has at most $\delta + 1$ coefficients, so the total size occupied by the output is $O(s\delta b)$.

5.5.1 Overview of the algorithm and its subroutines

5.5.1.1 over $\mathbb{Z}$

We start by presenting the main steps of the algorithm, leaving out some details of the analysis for the next subsection. Runtimes are given in terms of bit operations; here, we use the fact that operations $(+, \times)$ modulo a positive integer M take $O\tilde{~}(\log(M))$ bit operations, as does inversion modulo M if M is prime [73].

- **Introducing a change of coordinates.** We first choose a change of variables $\boldsymbol{\gamma}$ with coefficients in $\mathbb{Z}$. Applying it to the input equations $\mathcal{F}$ gives polynomials $\mathcal{H} = (h_1, \ldots, h_t)$, which we do not need to compute explicitly (as they may have large height). We let $\mathcal{B} = (B_0, \ldots, B_\sigma)$ be the lexicographic Gröbner basis of these polynomials in $\mathbb{Q}[x, y]$ (as with $\mathcal{H}$, we do not compute it explicitly).

  We assume that $\boldsymbol{\gamma}$ satisfies the assumptions of Propositions 5.4.1 to 5.4.3, so that their conclusions hold.

- **Computing Gröbner bases modulo $\boldsymbol{p}$.** Our second step is to choose two primes p, p', and compute the Gröbner bases $\mathcal{B}_p$ of $(\mathcal{H} \bmod p)$, and $\mathcal{B}_{p'}$ of $(\mathcal{H} \bmod p')$. We assume that neither p nor p' divides the integers $\beta_{\mathcal{F}}$ and $\beta_{\mathcal{H}}$ from Definition 29 applied to respectively $\mathcal{F}$ and $\mathcal{H}$. In particular, all denominators in $\mathcal{B}$ are invertible modulo p and p', and we have $\mathcal{B}_p = \mathcal{B} \bmod p$ and $\mathcal{B}_{p'} = \mathcal{B} \bmod p'$.

  To compute $\mathcal{B}_p$ and $\mathcal{B}_{p'}$, the algorithm reduces the $O(td^2)$ coefficients of $\mathcal{F}$ modulo p and p'. Then, we apply $\boldsymbol{\gamma}$ to the results, to obtain $\mathcal{H} \bmod p$ and $\mathcal{H} \bmod p'$. Due to Proposition 5.4.1, the coefficient of y^d in h_1 is a nonzero constant; if this is still the case modulo p and p', we use HermiteGroebnerBasis with $D = d$ to get $\mathcal{B}_p$ and $\mathcal{B}_{p'}$; otherwise, we raise an error.

Cost: Reducing the input coefficients take $\tilde{O}(td^2(h + \log(pp')))$ bit operations. Changing coordinates uses $\tilde{O}(td^2(\log(pp'))$ bit operations, by [73, Corollary 9.16]. Calling HermiteGroebnerBasis uses $\tilde{O}(t^\omega d^{\omega+1}(\log(pp')))$ bit operations, as we saw in Section 5.2.2.

- **Changing coordinates in $\mathcal{B}_p$ and $\mathcal{B}_{p'}$.** Using the Gröbner bases $\mathcal{B}_p$ and $\mathcal{B}_{p'}$ of $(\mathcal{H} \bmod p)$ and $(\mathcal{H} \bmod p')$, we compute the Gröbner bases of $(\mathcal{F} \bmod p)$ and $(\mathcal{F} \bmod p')$. This is done using the algorithm of [130]. Since pp' does not divide $\beta_{\mathcal{F}}$, we deduce that we obtain $\mathcal{G}_p = \mathcal{G} \bmod p$ and $\mathcal{G}_{p'} = \mathcal{G} \bmod p'$.

  Cost: This takes $\tilde{O}(\delta^3)$ operations in $\mathbb{F}_{p'}$, which is $\tilde{O}(\delta^3 \log(p'))$ bit operations.

- **Computing $\mathcal{G}_{p^k}$.** At each step of the main loop, we start from $\mathcal{G}_{p^{k/2}} = \mathcal{G} \bmod p^{k/2}$, and we compute $\mathcal{G}_{p^k} = \mathcal{G} \bmod p^k$. For this, we first need $\mathcal{F} \bmod p^k$; then, we use procedure LiftOneStepGroebner from [143, Remark 7.3] to obtain $\mathcal{G}_{p^k}$.

  Cost: Coefficient reduction takes $\tilde{O}(td^2(h + k\log(p)))$ bit operations. Algorithm LiftOneStepGroebner takes a one-time cost of $t\delta^\omega \log(p)$ bit operations, plus
  $$\tilde{O}(s^2 n_0 m_s + t\delta(d^2 + dm_s + s\delta))k\log(p))$$
  bit operations per iteration. Here, $n_0 = \deg_y(g_0)$ and $m_s = \deg_x(g_s)$.

- **Rational reconstruction.** We next attempt to recover all rational coefficients of $\mathcal{G}$ starting from those of $\mathcal{G}_{p^k} = \mathcal{G} \bmod p^k$. For each coefficient α of $\mathcal{G}_{p^k}$, we attempt to recover a pair (η, θ) in $\mathbb{Z} \times \mathbb{N}$, with $|\eta| < p^{k/2}/2$ and $\theta \le p^{k/2}$, θ invertible modulo p and $\alpha = \eta/\theta \bmod p^k$.

  Recall that we assume that all nonzero coefficients of $\mathcal{G}$ have numerators and denominators of height at most b. It follows that if $p^{k/2} > 2e^b$, we will succeed and correctly recover the corresponding coefficient in $\mathcal{G}$ [73, Theorem 5.26]. For smaller values of k, rational reconstruction may find no solution (in which case we reenter the lifting loop at precision $2k$), or may already terminate; in this case, its output $\mathcal{G}_{\mathrm{rec}}$ may be different from $\mathcal{G}$.

  Cost: Rational reconstruction takes $\tilde{O}(k\log(p))$ bit operations per coefficient, for a total of $\tilde{O}(s\delta k \log(p))$.

- **Testing for correctness.** The final step in the lifting loop is a randomized test, using $\mathcal{G}_{p'} = \mathcal{G} \bmod p'$ as a witness to detect those cases where rational reconstruction returned an incorrect result. We attempt to reduce $\mathcal{G}_{\mathrm{rec}}$ modulo our second prime p'; if this fails (because p' divides one of the denominators in it), we reenter the lifting loop at precision $2k$. Else, call $\mathcal{G}_{\mathrm{red}}$ the result. We simply compare $\mathcal{G}_{\mathrm{red}}$ and $\mathcal{G}_{p'} = \mathcal{G} \bmod p'$. If they coincide, we return $\mathcal{G}_{\mathrm{rec}}$, otherwise, we reenter the lifting loop.

Cost: Reduction modulo p' takes $O\tilde{}(b + \log(p'))$ bit operations per coefficient, for a total of $O\tilde{}(s\delta(b + \log(p'))$.

5.5.1.2 Over a function field

The method is mostly a copy of the case over $\mathbb{Z}$ up to some changes. The method of change of coordinates via $\mathrm{GL}_2(k[\boldsymbol{t}])$, calculating the Hermite normal form and the rational reconstruction and the test for correctness are a direct adaptation to $k[\boldsymbol{t}]$ with asymptotic complexity almost identical to the complexity over $\mathbb{Z}$ except for the $\log(p)$ factors which disappear as the operations are counted in terms of operations in k instead of bitwise operations in $\mathbb{F}_p$.

- **Introducing a change of coordinates.** Let $\mathcal{H} = (h_1, \ldots, h_t)$, with $\mathcal{G} = (b_0, \ldots, b_\sigma)$ its lexicographic Gröbner basis, be the image of $\mathcal{F}$ under a change of variables $\boldsymbol{\gamma}$ in $\mathrm{GL}_2(k[\boldsymbol{t}])$ (again we will not compute it explicitly). We assume that $\boldsymbol{\gamma}$ satisfies the assumptions of Propositions 5.4.1 and 5.4.3.

- **Computing Gröbner bases modulo a maximal ideal.** Our second step is equivalent to choose two elements t_0, t_1 in k, and compute the Gröbner bases of $\mathcal{H}$ at $\boldsymbol{t} = t_0$ ($\mathcal{B}_{\mathfrak{m}}$) and $\boldsymbol{t} = t_1$ ($\mathcal{B}_{\mathfrak{m}'}$).

  We assume that neither t_0 nor t_1 is a root of the polynomials $\beta_{\mathcal{F}}$ and $\beta_{\mathcal{H}}$ in $k[\boldsymbol{t}]$ from Definition 29. Supposing $|k|$ is sufficiently large, by the De Millo-Lipton-Schwartz-Zippel lemma, the probability that t_0 and t_1 are not root $\beta_{\mathcal{F}}$ and $\beta_{\mathcal{F}}$ is the at least $1/2^P$ if t_0, t_1 are sampled with with a uniform distribution in $\mathcal{S} \subset \mathbb{A}$ for $|\mathcal{S}| > 2^{P+1}(1 - C(t, d, D, h))$.

  Cost: Evaluating a t_0 and t_1 takes $O\tilde{}(td^2h)$ operations in k, where h is the maximal degree in $k[\mathbf{t}]$ of the coefficients in $\mathcal{F}$. Changing coordinates uses $O\tilde{}(td^2)$ operations. Calling HermiteGroebnerBasis uses $O\tilde{}(t^\omega d^{\omega+1})$ operations in k, as we saw in Section 5.2.2.

- **Changing coordinates in $\mathcal{B}_{\mathfrak{m}}$ and $\mathcal{B}_{\mathfrak{m}'}$.** Using the Gröbner basis $\mathcal{B}_{\mathfrak{m}}$ and $\mathcal{B}_{\mathfrak{m}'}$ of ($\mathcal{H} \bmod \langle t - t_0 \rangle$) and ($\mathcal{H} \bmod \langle t - t_1 \rangle$)), we compute the Gröbner basis $\mathcal{G}'$ of ($\mathcal{F} \bmod \langle \boldsymbol{t} - t_0 \rangle$) and ($\mathcal{F} \bmod \langle \boldsymbol{t} - t_1 \rangle$). Since $\beta_{\mathcal{F}}$ doe not vanish at t_1, we deduce that $\mathcal{G}_{\mathfrak{m}} = \mathcal{G} \bmod \langle \boldsymbol{t} - t_0 \rangle$ and $\mathcal{G}_{\mathfrak{m}'} = \mathcal{G} \bmod \langle \boldsymbol{t} - t_1 \rangle$.

Cost: This takes $O\tilde{~}(\delta^\omega)$ operations in k.

- **Computing $\mathcal{G}_{\mathfrak{m}^k}$.** At each step of the main loop, we start from $\mathcal{G}_{\mathfrak{m}^{k/2}} = \mathcal{G} \bmod \mathfrak{m}^{k/2}$, and we compute $\mathcal{G}_{\mathfrak{m}^k} = \mathcal{G} \bmod p^k$. For this, we first need $\mathcal{F} \bmod \mathfrak{m}^k$; then, we use procedure LiftOneStepGroebner from [143, Remark 7.3] to obtain $\mathcal{G}_{p^k}$.

  Cost: Coefficient reduction takes $O\tilde{~}(td^2(h+\kappa))$ operations in k, and changing coordinates $O\tilde{~}(td^2\kappa)$.

Using Remark 4.7.2 [143, Remark 7.4], and like we defined over $\mathbb{Z}$, the main cost, algorithm LiftOneStepGroebner, take

$$O\tilde{~}(t\delta^\omega + (\sigma^2 n_0 m_\sigma + t\delta(d^2 + dm_\sigma + \sigma\delta))\kappa)$$

operations in k.

Cost: Constructing the matrix and inverting it modulo $\langle \boldsymbol{t} - t_1\rangle^\kappa$ uses $O\tilde{~}(\delta^\omega\kappa)$ field operations.

- **Rational reconstruction.** To recover rational coefficients of $\mathcal{G}$ from those of $\mathcal{G}_{\mathfrak{m}^\kappa} = \mathcal{G} \bmod \langle \boldsymbol{t} - t_0\rangle^\kappa$. For each coefficient α of $\mathcal{G}_\kappa$, we attempt to recover a pair (η, θ) in $k[\boldsymbol{t}] \times k[\boldsymbol{t}]$, with $\deg \eta < \kappa/2$ and $\deg \theta \leq \kappa/2$, $\theta(t_0) \neq 0$ and $\alpha \cong \eta/\theta \mod \langle \boldsymbol{t} - t_0\rangle^\kappa$.

  Cost: Rational reconstruction takes $O\tilde{~}(\kappa)$ operations in k per coefficient, for a total of $O\tilde{~}(s\delta\kappa)$.

- **Testing for correctness.** The final step in the lifting loop is a randomized test, using $\mathcal{G}'$, to detect those cases where rational reconstruction returned an incorrect result. We attempt to reduce $\mathcal{G}_{\text{rec}}$ modulo our second prime $\langle \boldsymbol{t} - t_0\rangle$; if this fails (because $\langle \boldsymbol{t} - t_1\rangle$ divides one of the denominators in it), we reenter the lifting loop at precision 2κ. Else, call $\mathcal{G}_{\text{red}}$ the result. We simply compare $\mathcal{G}_{\text{red}}$ and $\mathcal{G}'$, which we know equals $\mathcal{G} \bmod \langle \boldsymbol{t} - t_0\rangle$. If they coincide, we return $\mathcal{G}_{\text{rec}}$. Otherwise, we reenter the lifting loop.

Cost: Evaluating at t_1 takes $O\tilde{~}(b)$ operations in k per coefficient, for a total of $O\tilde{~}(s\delta b)$.

5.5.2 Algorithm over general $\mathbb{A}$

Under the assumption there exists a fraction reconstruction algorithm (RationalReconstruction) and a valuation for $\mathbb{K}$, the following algorithm recovers the Gröbner basis of a zero-dimensional ideal.

Algorithm 5.5.1 GroebnerBasis($\mathcal{F}$)

```
Input: $\mathcal{F} = (f_1, \dots, f_t)$ in $\mathbb{A}[x, y]$
Output: the lexicographic Gröbner basis of $\mathcal{F}$ in $\mathbb{K}[x, y]$
choose two different maximal ideals $\mathfrak{m}, \mathfrak{m}'$ in $\mathbb{A}$
choose $\boldsymbol{\gamma}$ in $M_2(\mathbb{A})$
if $\boldsymbol{\gamma}$ mod $\mathfrak{m}$ or $\boldsymbol{\gamma}$ mod $\mathfrak{m}'$ is not invertible then raise an error
$\mathcal{H}_{\mathfrak{m}} \leftarrow$ ChangeCoordinates($\mathcal{F}, \boldsymbol{\gamma}$) mod $\mathfrak{m}$
$\mathcal{H}_{\mathfrak{m}'} \leftarrow$ ChangeCoordinates($\mathcal{F}, \boldsymbol{\gamma}$) mod $\mathfrak{m}'$
if the coefficient of $y^d$ in $h_1$ is zero then raise an error
$\mathcal{B}_{\mathfrak{m}} \leftarrow$ HermiteGroebnerBasis($\boldsymbol{h}, d$)
$\mathcal{B}_{\mathfrak{m}'} \leftarrow$ HermiteGroebnerBasis($\boldsymbol{h}', d$)
$\mathcal{G}_{\mathfrak{m}} \leftarrow$ ChangeCoordinatesGroebner($\mathcal{B}_{\mathfrak{m}}, \boldsymbol{\gamma}^{-1}$) mod $\mathfrak{m}$
$\mathcal{G}_{\mathfrak{m}'} \leftarrow$ ChangeCoordinatesGroebner($\mathcal{B}_{\mathfrak{m}'}, \boldsymbol{\gamma}^{-1}$) mod $\mathfrak{m}'$
$k \leftarrow 1$
repeat
    $k \leftarrow 2k$
    $\mathcal{G}_{\mathfrak{m}^k} \leftarrow$ LiftOneStepGroebner($\mathcal{F}$ mod $\mathfrak{m}^k, \mathcal{G}_{\mathfrak{m}^{k/2}}$)
    $b, \mathcal{G}_{\mathrm{rec}} \leftarrow$ RationalReconstruction($\mathcal{G}_{\mathfrak{m}^k}$)
    if not $b$ then continue
    $b, \mathcal{G}_{\mathrm{red}} \leftarrow \mathcal{G}_{\mathrm{rec}}$ mod $\mathfrak{m}'$
    if not $b$ then continue
until $\mathcal{G}_{\mathrm{red}} = \mathcal{G}_{\mathfrak{m}'}$
return $\mathcal{G}_{\mathrm{rec}}$
```

5.5.3 Parameters choice

5.5.3.1 over $\mathbb{Z}$

The change of variables $\boldsymbol{\gamma}$ needs to avoid a hypersurface $X \subset \overline{\mathbb{Q}}^4$ of degree at most $d^4 + d + \delta^3 + 3 \leq A_1 = d^6 + d^4 + d + 3$. We choose its entries uniformly at random in $\{0, \dots, 2^{P+2} A_1\}$; the cost of getting $\boldsymbol{\gamma}$ will be negligible.

Then, by the De Millo-Lipton-Schwartz-Zippel lemma, the probability that $\boldsymbol{\gamma}$ lies on X is at most $1/2^{P+2}$. In what follows, we assume that this is the case. As a result, all polynomials $\mathcal{H}$ have coefficients of height at most $h' = h + d(P + 5 + \log(A_1)) \in \tilde{O}(h + dP)$.

Let $\beta_{\mathcal{F}}$ and $\beta_{\mathcal{H}}$ be the nonzero integers from Definition 29 applied to respectively $\mathcal{F}$ and $\mathcal{H}$, and define

$$C_{\mathcal{F}} = C(t, d, \Delta_3(d), h) \in \tilde{O}(t^2 d^9 h) \quad \text{and} \quad C_{\mathcal{H}} = C(t, d, \Delta_1(d), h') \in \tilde{O}(t^2 d^4 hP).$$

Then, Proposition 5.3.1 gives upper bounds of the form $\text{height}(\beta_{\mathcal{F}}) \leq C_{\mathcal{F}}$ and $\text{height}(\beta_{\mathcal{H}}) \leq C_{\mathcal{H}}$. In particular, the height bound b on the coefficients of $\mathcal{G}$ satisfies $b \leq C_{\mathcal{F}}$, so b is in $\tilde{O}(t^2 d^9 h)$.

Let μ_1 be the coefficient of y^d in h_1, which has height at most h'. Our first requirement on p and p' is that neither of them divides $\mu = \beta_{\mathcal{F}} \beta_{\mathcal{H}} \mu_1$. This is a nonzero integer, with $\text{height}(\mu) \leq A_2$, where we set $A_2 = C_{\boldsymbol{f}} + C_{\mathcal{H}} + h' \in \tilde{O}(t^2 d^9 hP)$.

Finally, we want to ensure that in the verification step, if $\mathcal{G}_{\text{rec}}$ and $\mathcal{G}$ differ, their reductions modulo p', called $\mathcal{G}_{\text{red}}$ and $\mathcal{G}'$, differ as well. Below, we let κ_0 be the first power of two κ such that, at step κ, rational reconstruction correctly computes $\mathcal{G}_{\text{rec}} = \mathcal{G}$. For this, it suffices that $p^{\kappa/2} > 2e^b$, and one verifies this implies that $\kappa_0 \leq 8b \in \tilde{O}(t^2 d^9 h)$. Since all indices κ we go through are powers of two, there are at most $\log_2(8b)$ indices κ that could return an incorrect output.

Suppose then that at step $\kappa < \kappa_0$, we have found $\mathcal{G}_{\text{rec}}$ with rational coefficients; they all have numerators and denominators at most $p^{\kappa/2} \leq 2e^b$; on the other hand, the coefficients of $\mathcal{G}$ have numerators and denominators at most e^b. If $\mathcal{G}_{\text{rec}}$ and $\mathcal{G}$ differ, there exists a monomial whose coefficients in $\mathcal{G}_{\text{rec}}$ and $\mathcal{G}$ are different; it suffices to require that p' does not divide the numerator of their difference. This number has an absolute value of at most $4e^{2b}$.

Taking all $\kappa < \kappa_0$ into account, our last requirement is that p' also not divide a certain nonzero integer μ' (that depends on p). This integer μ' has height at most $\log_2(8b)(2b + \log(4))$, so that we have $\text{height}(\mu') \leq A_3$, with $A_3 = \log_2(8C_{\mathcal{F}})(2C_{\mathcal{F}} + \log(4)) \in \tilde{O}(t^2 d^9 h)$.

To summarize, once $\boldsymbol{\gamma}$ avoids X, it suffices that p does not divide μ and p' does not divide $\mu\mu'$ to ensure success. We can then finally make our procedure for choosing p and p' explicit:

- Let $B = 2^{P+3}\lceil A_2 \rceil$. We use the oracle $\mathscr{O}$ to obtain a uniformly sampled prime number in $[B + 1, \ldots, 2B]$. There are at least $B/(2\log(B))$ primes in this interval, and at most $\log(\mu)/\log(B)$ of them can divide μ, so the probability that p does is at most $2\log(\mu)/B$, which is at most $1/2^{P+2}$.

- Let $B' = 2^{P+3}\lceil A_2 + A_3 \rceil$. We use the oracle $\mathscr{O}$ to pick p' in the interval $[B' + 1, \ldots, 2B']$, and as a result, the probability that p' divides $\mu\mu'$ is at most $1/2^{P+2}$.

Altogether, the probability that $\boldsymbol{\gamma}$ avoids X, p does not divide μ and p' does not divide $\mu\mu'$ (and thus that the algorithm succeeds) is thus at least $1 - 3/2^{P+2} \geq 1 - 1/2^P$.

5.5.3.2 over $k[\boldsymbol{t}]$

The case over $k[\boldsymbol{t}]$ is simpler. Let $\mathcal{S} \subset k$, a change of variables $\boldsymbol{\gamma} \in \mathrm{GL}_2(\mathcal{S})$ avoid a hypersurface $X \subset \overline{k[\boldsymbol{t}]}^4$ of degree at most $d^4 + d + \delta^3 + 3 \leq \mathbb{A}_1 = d^6 + d^4 + d + 3$ with a probability less or equal to $\frac{A_1}{|\mathcal{S}|}$ by the De Millo-Lipton-Schwartz-Zippel lemma. Assuming $|k|$ is sufficiently large, chosing $|\mathcal{S}| \geq 2^P(A_1)$, we have the same probability bound as defined over $\mathbb{Z}$. Since $|\ |$ is ultrametric, the degree bound for the coefficients in $\mathcal{H}$ is the same as $\mathcal{F}$. For the degree of the coefficients in $\mathcal{G}$, b, we still rely on Proposition 5.3.1.

5.5.4 Runtime analysis

5.5.4.1 over $\mathbb{Z}$

We assume that choosing a random integer in a set $\{0, \ldots, A\}$ (with the uniform distribution) uses $O\tilde{}(\log(A))$ bit operations. Since we do not want to discuss algorithms for prime generation, we also assume that have an oracle $\mathscr{O}$, which takes as input an integer C, and returns a prime number in $I = [C+1, \ldots, 2C]$, uniformly distributed within the set of primes in I, using $O\tilde{}(\log(C))$ bit operations.

To give our final runtime estimate over $\mathbb{Z}$, we first note that both $\log(p)$ and $\log(p')$ are in $O\tilde{}(P + \log(tdh))$. Besides, the definition of κ_0 implies that at all lifting steps, $\kappa \log(p)$ is in $O\tilde{}(b + \log(p))$, that is $O\tilde{}(b + P + \log(tdh))$. After some straightforward simplifications, the runtime becomes the sum of the following terms

- $O\tilde{}(td^2h)$
- $O\tilde{}((t^\omega d^{\omega+1} + \delta^\omega)(P + \log(tdh))$
- $O\tilde{}((t\delta(d^2 + dm_\sigma + \sigma\delta) + \delta^\omega)(b + P + \log(tdh)))$.

In order to get a better grasp on this runtime, let us assume that P is a fixed constant, and use the upper bound $\sigma \leq m_\sigma \leq \delta$. This yields the overall bound

$$O\tilde{}(td^2h + (t^\omega d^{\omega+1} + \delta^\omega)\log(tdh) + (td^2\delta + t\delta^3)(b + \log(tdh))),$$

where we recall that the input size is $O(td^2h)$ bits, and the output size $O(s\delta b) \subset O(\delta^2 b)$ bits. The $\log(tdh)$ factors can be omitted from the resulting complexity as they are dominated by higher terms (see detail at the end of this chapter).

In order to get a better grasp on this runtime, let us assume that P and the number of equations t are fixed constants, and use the upper bounds $n_0, m_s \leq \delta$. This gives a total bound softly linear in

$$d^2h + (d^{\omega+1} + \delta^\omega)\log(h) + (d^2\delta + d\delta^2 + s^2\delta^2)(b + \log(h)).$$

The first term is the input size, the second one describes computations done modulo small primes, and the last one computations done modulo higher powers of p. We also recall that the output size $O(s\delta b)$ bits.

5.5.4.2 over $k[\boldsymbol{t}]$

The case over $k[\boldsymbol{t}]$ is simpler. We now assume that choosing a random element with the uniform distribution in $\mathcal{S} \subseteq k$ can be done in $O(log(|\mathcal{S}|))$.

Here the running time is direct and similar mostly similar to the results obtained over $\mathbb{Z}$ except for the $log(p)$, which do not appear from the beginning; in this case, the runtime is expressed in terms of operations in the base field k:

- $O\tilde{}(td^2h)$
- $O\tilde{}((t^\omega d^{\omega+1} + \delta^\omega)(P + \log(d^6 + d + 3))$
- $O\tilde{}((t\delta(d^2 + dm_\sigma + \sigma\delta) + \delta^\omega)(b + P))$.

If we fix P to a constant, it simplifies to

$$O\tilde{}(td^2h + (t^\omega d^{\omega+1} + \delta^\omega) + (td^2\delta + t\delta^3)b).$$

Note that for a fixed probability P, the cost over $k[\boldsymbol{t}]$ (in k operation) is the same as over $\mathbb{Z}$ (in bit operation) up to a logarithmic factor.

5.5.5 Variant

We describe here how to modify the algorithm over $\mathbb{Q}$ if we are only interested in the Gröbner basis $\mathcal{G}^0 = (g_0^0, \dots, g_r^0)$ of the $\langle x, y\rangle$-primary component of I.The algorithm can be directly adapted to other fields that meet our requirements (endowed with a fraction reconstruction algorithm, large characteristic or zero and a valuation), *e.g.* $\Bbbk(t)$. In what follows, we let η be the degree of this ideal, and c be the maximum height of the numerators and denominators of the coefficients of $\mathcal{G}^0$. Hence, the input has total size $O(td^2h)$, and the output $O(r\eta c)$.

As above, we use a change of coordinates $\boldsymbol{\gamma}$, and we call $\mathcal{B}^0 = (B_0^0, \dots, B_\rho^0)$ the Gröbner basis of the $\langle x, y\rangle$-primary component of $I^{\boldsymbol{\gamma}}$.

- The first difference is that we now use GroebnerBasisAtZero instead of HermiteGroebnerBasis, modulo p and p'. Since we are in generic coordinates, we can use degree $D = d$, so the runtime is $O\tilde{}(td^{\omega}m_{\rho}(\log(p) + \log(p')))$ bit operations, where m_{ρ} is the x-degree of B^0_{ρ}.
- The lifting itself is done using the algorithm LiftOneStepPunctualGroebnerBasis from Remark 4.7.1 [143, Remark 7.3]. This time, the cost is $O\tilde{}(t\eta^{\omega}\log(p) + t\eta^2 m_{\rho}\kappa\log(p))$ bit operations.

The rest of the algorithm is unchanged. The conditions that guarantee success are slightly different as well.

- Now, $\boldsymbol{\gamma}$ must avoid a hypersurface Y of degree at most $d + d^4 + \eta^3 + 3 \leq d^6 + d^4 + d + 3$, in order to guarantee that $I^{\boldsymbol{\gamma}}$ satisfies Propositions 5.4.1 to 5.4.3.
- The primes p and p' should divide the denominator of no coefficient in the Gröbner bases $\mathcal{G}^0$ and $\mathcal{B}^0$; besides, these polynomials reduced modulo p (resp. p') should still define the $\langle x, y\rangle$-primary components of $f_1 \bmod p, \ldots, f_t \bmod p$ and $f_1^{\boldsymbol{\gamma}} \bmod p, \ldots, f_t^{\boldsymbol{\gamma}} \bmod p$ (resp. modulo p').

  We use the fact that the $\langle x, y\rangle$-primary component of $\langle f_1, \ldots, f_t\rangle$ is the ideal generated by $\mathcal{F} = (f_1, \ldots, f_t, x^{d^2}, y^{d^2})$; similarly for $\mathcal{H} = (f_1^{\boldsymbol{\gamma}}, \ldots, f_t^{\boldsymbol{\gamma}})$, giving us polynomials $\boldsymbol{H}$. It is then sufficient that neither p nor p' divides the integers $\beta_{\mathcal{F}}\beta_{\boldsymbol{H}}$ from Definition 29. Their heights are in $O\tilde{}(t^2d^6h)$ and $O\tilde{}(t^2d^6h')$, where h' is the height bound on $\mathcal{H}$.

The rest of the analysis is conducted as before. Given a fixed integer P, we deduce that we can compute the Gröbner basis $\mathcal{G}^0$, with probability of success at least $1 - 1/2^P$, using $O\tilde{}(td^2h + (td^{\omega}\eta + \eta^{\omega})\log(tdh) + t\eta^3(c + \log(tdh)))$ binary operations. Once again, the $\log(tdh)$ factors can disappear from the resulting complexity as they are dominated by polynomial expressions in t, d and h.

The section below presents why the logarithmic factors in Section 5.5.4 and Section 5.5.5 can be omitted.

Proposition 5.5.1. $O\tilde{}(td^2h + (t^\omega d^{\omega+1} + \delta^\omega)\log(tdh) + (td^2\delta + t\delta^3)(b + \log(tdh))) = O\tilde{}(td^2h + (t^\omega d^{\omega+1} + \delta^\omega) + (td^2\delta + t\delta^3)b)$.

Proof. First, we observe that

$$td^2h + (t^\omega d^{\omega+1} + \delta^\omega)\log(tdh) + (td^2\delta + t\delta^3)(b + \log(tdh))$$

is smaller than

$$(td^2h + (t^\omega d^{\omega+1} + \delta^\omega) + (td^2\delta + t\delta^3)b)\log(tdh).$$

Let $F = (td^2h)$, since $t \geq 0, d \geq 0$ and $h \geq 0$ then $F \geq t \implies \log(t) \leq \log(F)$. The same argument holds for d and h. Thus $\log(tdh) \leq \log(F)$ thus they are omitted in $O\tilde{}(F)$ notation. □

By the same argument one can show that $O\tilde{}(td^2h + (td^\omega\eta + \eta^\omega)\log(tdh) + t\eta^3(c + \log(tdh))) = O\tilde{}(td^2h + (td^\omega\eta + \eta^\omega) + t\eta^3c)$.

Chapter 6

Discussion

Our main results are statements regarding Gröbner bases of zero-dimensional ideals or their primary component(s); we complete the analysis for the $\langle x, y\rangle$-primary component. We also explore the possibility of changing bases, via the map tangling and untangling, to use our result for other primary components. Still, the complete analysis for general primary components is left as future work.

Let $I = \bigcap_{Q\in\mathcal{Q}} Q$ be a primary decomposition of an ideal $I \subseteq \mathbb{K}[x, y]$. Through the Chinese remainder theorem

$$K[x, y]/I \cong \prod K[x, y]/ \bigcap_{Q\in\mathcal{Q}} Q$$

both representation, *i.e.* the Gröbner basis of I or the Gröbner bases of all $Q \in \mathcal{Q}$, are equivalent. Although the complexity of passing from one to the other is not evaluated here. Nonetheless, for our goal, the latest (a basis for $Q \in \mathcal{Q}$) is more convenient as it highlights the local structure of a point (we are already localised), particularly when the component is centred at the origin.

In general, we expect the multiplicity of each intersection to be significantly smaller than the degree of the ideal. Hence, overall, finding all the bases of primary components separately should take less time than the same operation for the complete ideal since the complexity is more than linear in terms of the degree.

To adapt the result of Section 5.5.5 to any primary component, we would need to complete the complexity analysis over a base ring of the form $\mathbb{A}[\alpha_1, \ldots, \alpha_r]$, where $\alpha_1, \ldots, \alpha_r$ are algebraic in $\mathbb{A}$. If a fraction reconstruction algorithm exists for $\mathbb{K}$ the fraction field of a ring $\mathbb{A}$, the fraction reconstruction over $\mathbb{K}[\alpha_1, \ldots, \alpha_r]$ could be performed by linear algebra. Further, there could be some alterations to the basis between the tangled and untangled components, such as some reductions, to ensure the basis is still a reduced minimal Gröbner basis after the change of basis. This is a challenge we will address soon.

6.0.1 Note on Newton iteration and field extension

In Section 2.1.6, we stated that Newton's method could be applied to algebraic values. This is simple in the case where we know in which extension the root lives, which is rarely the case. Luckily, this is the case for our coefficients in Theorem 4.1.1.

However, in general, for points belonging to an unknown extension, it usually further requires finding the polynomial(s) that defines the extension. In those situations, a Newton iteration for Gröbner basis as we defined in this thesis or the Hermite/Howell normal form approach would be efficient to define the extension to be used.

6.0.2 The base ring

The complete algorithm, GroebnerBasis presented in Section 5.5, works over any field $\mathbb{K}$ with a valuation (Archimedean or ultrametric) and for which there exists a fraction reconstruction algorithm with respect to the $\mathfrak{m}$-adic filtration for $\mathfrak{m} \subseteq \mathbb{A}$ a maximal ideal over a domain $\mathbb{A} \subset \mathbb{K}$. We walk through the details over $\mathbb{A} = \mathbb{Z}$ ($\mathbb{K} = \mathbb{Q}$) and $\mathbb{A} = k[\boldsymbol{t}]$ ($\mathbb{K} = k(\boldsymbol{t})$), to provide complete examples for the two types of valuations . The Newton iterator, LiftOneStep from Theorem 4.1.1, has no constraint for $\mathbb{K}$. The validation and the fraction reconstruction algorithm are required to recover an exact expression in $\mathbb{K}$ from a partial representation in the $\mathfrak{m}$-adic completion. Having a field of large characteristics allows us to benefit from the Borel-fixed property obtained through the change of basis, which is valuable for complexity purposes but not essential to the algorithm. For some base rings, such as $k(\boldsymbol{t})$ where the operations are counted in the field k, then the larger charateristique is further required to ensure the existence of an element in k which is not a root of a given polynomial in $k(\boldsymbol{t})$. We highlight that if $\mathbb{A}$ and $\mathbb{K}$ satisfy the above condition, so does their algebraic extension. Thus the algorithm should be adaptable to the image ring obtained by tangling.

6.0.3 Borel-fixed

The statement in Section 5.4 that the initial term ideals of zero-dimensional ideals are Borel-fixed in generic coordinate can be deduced from [144], which proves the statement for any weighted ordering. The weighted orderings can emulate any the other ordering, *e.g.* the weighted ordering with

$$((1, 0, 0, \dots), (0, 1, 0, \dots), \dots, (0, 0, \dots, 1))$$

is equivalent to the lexicographic ordering [45, §2.4][61, proposition 15.16]. The reason why the approach is interesting here is purely to characterize the Zariski open of $\mathrm{GL}_2(\mathbb{K})$ to establish the probability that a change of basis leads to the Borel-fixed property that we need for our analysis.

6.1 A few open questions

to my future self,

In this document, we focus only on the lexicographic order which raises the following question.

Question 6.1.1. *What about other monomial orders?*

While the lexicographic ordering is sufficient to describe the local structure, some applications are better performed using different ordering (see Appendix C for examples), so the question is relevant. However, since an explicit bijection between moduli space and a stratum is only known for the lexicographic order, we do not discuss any other monomial ordering. There are, however, algorithms such as FGLM, introduced **F**augère, **G**ianni, **L**azard and **M**ora in 1993 [68], that convert a base for a given monomial order to a distinct order. Thence, one can apply the current algorithm and then employ FGLM to a distinct monomial ordering. Analysis for this procedure falls outside the scope of the current thesis and is left as future work. It is unclear if it would compare favourably to other approaches.

Still, it would be interesting to define an explicit map between the moduli space of a Gröbner cell through the syzygies of the generators in a Gröbner basis in different orderings. Even more, since the coefficients obtained for lexicographic ordering are empirically larger than other orderings, *e.g.* graded orderings, so using a different ordering could lead to a better complexity. Adapting our limit construction algorithm to a different ordering would also entail rethinking the efficient modulo operations with a basis in a different ordering and the results that follow from the Hermite normal form.

Question 6.1.2. *Could the method be adapted for n-dimensional curves?*

It is well known that the moduli spaces are also affine spaces. Again, it would be interesting to define an explicit map between the moduli space of a stratum and its ideals for more than two variables. It might be feasible, although such a map is not currently known to us. I keep the question open for the moment, but I would love to explore it in a near future.

6.1.1 Bound on the coefficients

It would be interesting to get a sharper bound on the growth of the coefficients in a lexicographic Gröbner basis. Other than the approach using Hermite, we also explored the avenue of adapting the arithmetic Nullstellensatz results of Krick, Pardo, and Sombra [105] to use it on the ideal of coefficients defined in Section 4.6.2 (see Corollary 4.6.1 and Corollary 4.6.2). We looked into the reduction of monomials in the support of an ideal I by a parametric Gröbner basis of a stratum to the reduction into blocks that use different sets of parameters. However, the analysis did not lead to a more optimal bound: the bound obtained was less sharp than the one obtained with the Hermite normal form, so this approach was omitted from this thesis. My intuition is that additional structure should be taken into account. It would be interesting to see what can be done when looking at the reduction of the elements of I instead of working on the supports.

6.1.2 Primary decomposition

Modular operations with $\langle x, y\rangle$-primary components are faster and the Hilbert-Burch matrix of punctual basis requires fewer variables. Whence, complexity-wise, it is more interesting to move a component at the origin when working with a primary ideal; thus, the idea of using a morphism like tangling and untangling. We would be interested in the complexity of doing the operation of change of basis without hypotheses on the primary ideal.

Question 6.1.3. *What is the arithmetic complexity of the map tangling and untangling for primary components that are not monomial when seen through untangling?*

It should not be too complicated to lift the hypothesis. This would bring us a step forward toward an algorithm to find the primary decomposition of a zero-dimensional bivariate ideal I written as an intersection of Gröbner basis. In particular, assume $\mathbb{K}$ meets our conditions (large field, endowed with a fraction reconstruction algorithm, fraction field of a domain $\mathbb{A}$) and let $\mathcal{F} \subset \mathbb{A}[x, y]$be a finite subset and $J = \langle \mathcal{F} \rangle \subseteq \mathbb{K}[x, y]$ an ideal, we could:

- apply the method like [124] to find $V(I)$;
- if $V(I)$ is represented as the shape lemma $(u(x), v(x))$, factorize $u(x)$, which is square-free, into irreducibles $u(x) = \prod_{w \in \mathcal{W}} w$ then for all factors $w \in \mathcal{W}$ we can simplify simplify $v(x)$, denoted $v \mid_w$;
- for all irreducible $w \in \mathcal{W}$, if $p \in V(w, v \mid_w) \subset V(I)$ has multiplicity greater than 1, use the result for untangling to move a primary component at the origin in an alternative base ring $\mathbb{A}[\alpha_1, \ldots, \alpha_r]$ for α_i algebraic;
- apply the method of Howell normal form to find the Gröbner basis of the primary component at the origin in a bivariate polynomial ring over $\mathbb{A}[\alpha_1, \ldots, \alpha_r]/\mathfrak{m}$ for $\mathfrak{m} \subseteq \mathbb{A}$ maximal;
- use the Newton iterator to find the Gröbner to lift the basis and use linear system solving to do a rational reconstruction in $\mathbb{K}[\alpha_1, \ldots, \alpha_r]$;

To use genericity assumptions, *e.g.* the Borel-fixed property, some changes of coordinates may be required in the above road map. From there, we would already get the local structure of all the points. To get the primary decomposition, we shall further return to the original ring:

- applies a change on the basis (tangling) and possibly a change of coordinates.
- Additional steps might be needed, *e.g.* reduction might be needed to have a reduce minimal Gröbner basis.

But let us keep that for a next adventure.

References

[1] W. A. Adkins and S. H. Weintraub. *Algebra: an approach via module theory*, volume 136. Springer Science & Business Media, 2012.

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

[3] L. Alberti, B. Mourrain, and J. Wintz. Topology and arrangement computation of semi-algebraic planar curves. *Computer Aided Geometric Design*, 25(8):631–651, 2008.

[4] L. Alberti, B. Mourrain, and J. Wintz. Topology and arrangement computation of semi-algebraic planar curves. *Computer Aided Geometric Design*, 25(8):631–651, 2008.

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

[6] E. A. Arnold. Modular algorithms for computing Gröbner bases. *Journal of Symbolic Computation*, 35(4):403–419, 2003.

[7] M. Atiyah. *Introduction to commutative algebra.* CRC Press, 2018.

[8] P. Aubry, D. Lazard, and M. Moreno Maza. On the theories of triangular sets. *Journal of Symbolic Computation*, 28(1,2):45–124, 1999.

[9] C. W. Ayoub. On constructing bases for ideals in polynomial rings over the integers. *Journal of Number Theory*, 17(2):204–225, 1983.

[10] D. J Bates, A. J Sommese, J. D Hauenstein, and C. W Wampler. *Numerically solving polynomial systems with Bertini.* SIAM, 2013.

[11] D. Bayer and M. Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Mathematical Journal*, 55(2):321–328, 1987.

[12] D. Bayer and M. Stillman. Computation of Hilbert functions. *Journal of Symbolic Computation*, 14(1):31–50, 1992.

[13] K. Belabas, M. van Hoeij, J. Klüners, and A. Steel. Factoring polynomials over global fields. *Journal de théorie des nombres de Bordeaux*, 21(1):15–39, 2009.

[14] E. Berberich, P. Emeliyanenko, and M. Sagraloff. An elimination method for solving bivariate polynomial systems: Eliminating the usual drawbacks. In *ALENEX*, pages 35–47. SIAM, 2011.

[15] G. M Bergman. The diamond lemma for ring theory, 2008.

[16] J. Berthomieu, B. Boyer, and J.-C. Faugère. Linear algebra for computing gröbner bases of linear recursive multidimensional sequences. *Journal of Symbolic Computation*, 83:36–67, 2017.

[17] J. Berthomieu and R. Lebreton. Relaxed p-adic Hensel lifting for algebraic systems. In *ISSAC'12*, pages 59–66. ACM, 2012.

[18] L. A Bokut and Y. Chen. Gröbner–Shirshov bases and their calculation. *Bulletin of Mathematical Sciences*, 4(3):325–395, 2014.

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

[20] A. Bostan, F. Chyzak, M. Giusti, R. Lebreton, G. Lecerf, B. Salvy, and É. Schost. *Algorithmes Efficaces en Calcul Formel*. Palaiseau, September 2017. 686 pages. Imprimé par CreateSpace. Aussi disponible en version électronique.

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

[22] A. Bostan, C.-P. Jeannerod, and É. Schost. Solving structured linear systems with large displacement rank. *Theor. Comput. Sci.*, 407(1-3):155–181, 2008.

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

[24] Y. Bouzidi. *Résolution de systèmes bivariés et topologie de courbes planes*. PhD thesis, Université de Lorraine, 2014.

[25] Y. Bouzidi, S. Lazard, G. Moroz, M. Pouget, and F. Rouillier. Improved algorithm for computing separating linear forms for bivariate systems. In *ISSAC'14*, pages 75–82. ACM, 2014.

[26] Y. Bouzidi, S. Lazard, G. Moroz, M. Pouget, F. Rouillier, and M. Sagraloff. Solving bivariate systems using rational univariate representations. *Journal of Complexity*, 37:34–75, 2016.

[27] Y. Bouzidi, S. Lazard, M. Pouget, and F. Rouillier. Rational univariate representations of bivariate systems and applications. In *ISSAC'13*, pages 109–116. ACM, 2013.

[28] Y. Bouzidi, S. Lazard, M. Pouget, and F. Rouillier. Separating linear forms for bivariate systems. In *ISSAC'13*, pages 117–124. ACM, 2013.

[29] G. E. Bredon. *Topology and geometry*, volume 139. Springer Science & Business Media, 2013.

[30] M. R Bremner and V. Dotsenko. *Algebraic operads: an algorithmic companion*. CRC Press, 2016.

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

[32] J. Briançon. Description de $\mathrm{Hilb}^n\ \mathbb{C}\{x,y\}$. *Inventiones Mathematicae*, 41:45–90, 1977.

[33] J. Briançon and A. Galligo. Déformations distinguées d'un point de $\mathbb{C}^2$ ou $\mathbb{R}^2$. In *Singularités à Cargèse*, number 7-8 in Astérisque, pages 129–138. Société mathématique de France, 1973.

[34] B. Buchberger. A note on the complexity of constructing gröbner-bases. In *European Conference on Computer Algebra*, pages 137–145. Springer, 1983.

[35] J. Canny. Generalised characteristic polynomials. *Journal of Symbolic Computation*, 9:241–250, 1990.

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

[37] D. G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica*, 28(7):693–701, 1991.

[38] J. Cheng, S. Lazard, L. M. Peñaranda, M. Pouget, F. Rouillier, and E. P. Tsigaridas. On the topology of real algebraic plane curves. *Mathematics in Computer Science*, 4(1):113–137, 2010.

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

[40] A. Conca and G. Valla. Canonical Hilbert-Burch matrices for ideals of $k[x, y]$. *Michigan Mathematical Journal*, 57:157 – 172, 2008.

[41] A. Conca and M. Varbaro. Square-free gröbner degenerations. *Inventiones mathematicae*, 221(3):713–730, 2020.

[42] A. Constantinescu. Parametrizations of ideals in $k[x, y]$ and $k[x, y, z]$. *Journal of Algebra*, 346(1):1–30, 2011.

[43] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990.

[44] D. A Cox, J. Little, and D. O'Shea. *Using algebraic geometry*, volume 185. Springer Science & Business Media, 2006.

[45] D. A Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*. Springer Science & Business Media, 2013.

[46] X. Dahan. Size of coefficients of lexicographical Groöbner bases: The zero-dimensional, radical and bivariate case. In *ISSAC'09*, page 119–126. ACM, 2009.

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

[48] X. Dahan. Lexicographic Gröbner bases of bivariate polynomials modulo a univariate one. *Journal of Symbolic Computation*, 110:24–65, 2022.

[49] X. Dahan, M. Moreno Maza, É. Schost, W. Wu, and Y. Xie. Lifting techniques for triangular decomposition. In *ISSAC'05*. ACM press, 2005.

[50] X. Dahan and É. Schost. Sharp estimates for triangular sets. In *ISSAC'04*, pages 103–110. ACM, 2004.

[51] B. Dayton, T.-Y. Li, and Z. Zeng. Multiple zeros of nonlinear systems. *Mathematics of Computation*, 80(276):2143–2168, 2011.

[52] L. De Feo and É. Schost. Fast arithmetics in artin-schreier towers over finite fields. *Journal of Symbolic Computation*, 47(7):771–792, 2012.

[53] W. Decker, G.-M. Greuel, and G. Pfister. Primary decomposition: algorithms and comparisons. In *Algorithmic Algebra and Number Theory: Selected Papers From a Conference Held at the University of Heidelberg in October 1997*, pages 187–220. Springer, 1999.

[54] D. N. Diatta, S. Diatta, F. Rouillier, M.-F. Roy, and M. Sagraloff. Bounds for polynomials on algebraic numbers and application to curve topology, 2021.

[55] A. Dickenstein, N. Fitchas, M. Giusti, and C. Sessa. The membership problem for unmixed polynomial ideals is solvable in single exponential time. *Discrete Applied Mathematics*, 33(1):73–94, 1991.

[56] D. I. Diochnos, I. Z. Emiris, and E. P. Tsigaridas. On the asymptotic and practical complexity of solving bivariate systems over the reals. *Journal of Symbolic Computation*, 44(7):818–835, 2009.

[57] R. Duan, H. Wu, and R. Zhou. Faster matrix multiplication via asymmetric hashing. *arXiv preprint arXiv:2210.10173*, 2022.

[58] D. S. Dummit and R. M Foote. *Abstract algebra*, volume 3. Wiley Hoboken, 2004.

[59] G. L. Ebert. Some comments on the modular approach to Gröbner bases. *ACM SIGSAM Bull.*, 17(2):28–32, 1983.

[60] S. Eilenberg and H. P. Cartan. *Homological algebra.* Princeton University Press, 1956.

[61] D. Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[62] D. Eisenbud and J. Harris. *The geometry of schemes*, volume 197. Springer Science & Business Media, 2006.

[63] D. Eisenbud, H. I Levine, and B. Teissier. An algebraic formula for the degree of a c^∞ map germ/sur une inégalité à la minkowski pour les multiplicités. *Annals of Mathematics*, pages 19–44, 1977.

[64] M. El Kahoui. Topology of real algebraic space curves. *Journal of Symbolic Computation*, 43(4):235–258, 2008.

[65] G. Ellingsrud and S. Strø mme. On the homology of the Hilbert scheme of points in the plane. *Inventiones Mathematicae*, 87:343–352, 1987.

[66] P. Emeliyanenko and M. Sagraloff. On the complexity of solving a bivariate polynomial system. In *ISSAC'12*, pages 154–161. ACM, 2012.

[67] I. Z. Emiris and E. P. Tsigaridas. Real solving of bivariate polynomial systems. In *CASC*, pages 150–161. Springer, 2005.

[68] J.-C. Faugere, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional gröbner bases by change of ordering. *Journal of Symbolic Computation*, 16(4):329–344, 1993.

[69] G. Carrà Ferro. Gröbner bases and Hilbert schemes. i. *Journal of Symbolic Computation*, 6(2):219–230, 1988.

[70] C. M. Fiduccia. An efficient formula for linear recurrences. *SIAM Journal on Computing*, 14(1):106–112, 1985.

[71] A. Friedman and B. Hu. Bifurcation from stability to instability for a free boundary problem arising in a tumor model. *Archive for rational mechanics and analysis*, 180:293–330, 2006.

[72] A. Galligo. A propos du théoreme de préparation de Weierstrass. In *Fonctions de plusieurs variables complexes*, pages 543–579. Springer, 1974.

[73] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, third edition, 2013.

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

[75] J. Gerhard. Modular algorithms for polynomial basis conversion and greatest factorial factorization. In *RWCA'00*, pages 125–141, 2000.

[76] P. Gianni and T. Mora. Algebraic solution of systems of polynomial equations using Groebner bases. In *AAECC'87*, volume 356 of *Lecture Notes in Comput. Sci.*, pages 247–257. Springer, 1989.

[77] M. Giesbrecht. Nearly optimal algorithms for canonical matrix forms. *SIAM Journal on Computing*, 24(5):948–969, 1995.

[78] H. Gillet and C. Soulé. Arithmetic intersection theory. *Publications Mathématiques de l'IHÉS*, 72:93–174, 1990.

[79] M. Giusti, K. Hägele, J. Heintz, J.-E. Morais, J.-L. Montaña, and L.-M. Pardo. Lower bounds for diophantine approximation. *J. of Pure and Applied Algebra*, 117/118:277–317, 1997.

[80] M. Giusti and J. Heintz. 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*, volume XXXIV, pages 216–256, 1993.

[81] M. Giusti, J. Heintz, J.-E. Morais, J. Morgenstern, and L.-M. Pardo. Straight-line programs in geometric elimination theory. *Journal of Pure and Applied Algebra*, 124:101–146, 1998.

[82] M. Giusti, J. Heintz, J.-E. Morais, and L.-M. Pardo. When polynomial equation systems can be solved fast? In *AAECC-11*, volume 948 of *LNCS*, pages 205–231. Springer, 1995.

[83] M. Giusti, G. Lecerf, and B. Salvy. A Gröbner-free alternative for polynomial system solving. *Journal of Complexity*, 17(1):154–211, 2001.

[84] L. González-Vega and M. El Kahoui. An improved upper complexity bound for the topology computation of a real algebraic plane curve. *Journal of Complexity*, 12(4):527 – 544, 1996.

[85] J. Harris. *Algebraic geometry: a first course*, volume 133. Springer Science & Business Media, 2013.

[86] R. Hartshorne. *Algebraic geometry*, volume 52. Springer Science & Business Media, 2013.

[87] A. Hashemi and D. Lazard. Sharper complexity bounds for zero-dimensional gröbner bases and polynomial system solving. *International Journal of Algebra and Computation*, 21(05):703–713, 2011.

[88] J. D. Hauenstein, B. Mourrain, and A. Szanto. On deflation and multiplicity structure. *Journal of Symbolic Computation*, 83:228–253, 2017. Special issue on the conference ISSAC 2015: Symbolic computation and computer algebra.

[89] J. Heintz and C. P. Schnorr. Testing polynomials which are easy to compute (extended abstract). In *STOC*, pages 262–272. ACM, 1980.

[90] J. van der Hoeven and R. Larrieu. Fast reduction of bivariate polynomials with respect to sufficiently regular Gröbner bases. In *ISSAC '18*, page 199–206. ACM Press, 2018.

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

[92] R. Homs and A.-L. Winz. Canonical Hilbert-Burch matrices for power series. *Journal of Algebra*, 583:1–24, 2021.

[93] J. A. Howell. Spans in the module $\mathbb{Z}_m^s$. *Linear and Multilinear Algebra*, 19(1):67–77, 1986.

[94] S. G. Hyun, S. Melczer, É. Schost, and C. St-Pierre. Change of basis for $\mathfrak{m}$-primary ideals in one and two variables. In *ISSAC'19*, pages 227–234. ACM Press, 2019.

[95] A. Iarrobino. *Punctual Hilbert Schemes*. Number 188 in Memoirs of the American Mathematical Society. American Mathematical Society, 1977.

[96] O. H. Ibarra, S. Moran, and R. Hui. A generalization of the fast LUP matrix decomposition algorithm and applications. *J. Algorithms*, 3(1):45–56, 1982.

[97] J. van der Hoeven and É. Schost. Multi-point evaluation in higher dimensions. *Applicable Algebra in Engineering, Communication and Computing*, 24(1):37–52, 2013.

[98] N. Jacobson. Structure theory for algebraic algebras of bounded degree. *Annals of Mathematics*, pages 695–707, 1945.

[99] C.-P. Jeannerod. LSP matrix decomposition revisited. 2006.

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

[101] K. S. Kedlaya and C. Umans. Fast polynomial factorization and modular composition. *SIAM Journal on Computing*, 40(6):1767–1802, 2011.

[102] A. Kobel and M. Sagraloff. Improved complexity bounds for computing with planar algebraic curves. *CoRR*, abs/1401.5690, 2014.

[103] A. Kobel and M. Sagraloff. On the complexity of computing with planar algebraic curves. *Journal of Complexity*, 31(2):206–236, 2015.

[104] J. Kollar. Sharp effective nullstellensatz. *Journal of the American Mathematical Society*, 1(4):963–975, 1988.

[105] T. Krick, L. M. Pardo, and M. Sombra. Sharp estimates for the arithmetic nullstellensatz. 2001.

[106] G. Labahn, V. Neiger, T. X. Vu, and W. Zhou. Rank-sensitive computation of the rank profile of a polynomial matrix. In *ISSAC'22*, page 351–360. ACM, 2022.

[107] G. Labahn, V. Neiger, and W. Zhou. Fast, deterministic computation of the Hermite normal form and determinant of a polynomial matrix. *Journal of Complexity*, 42:44–71, 2017.

[108] S. Lang. *Algebra*, volume 211. Springer Science & Business Media, 2012.

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

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

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

[112] G. Lecerf. Quadratic Newton iteration for systems with multiplicity. *Foundations of Computational Mathematics*, 2(3):247–293, 2002.

[113] P. Lella and M. Roggero. Rational components of Hilbert schemes. *Rend. Semin. Mat. Univ. Padova*, 126:11–45, 2011.

[114] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische annalen*, 261(ARTICLE):515–534, 1982.

[115] A. Leykin, J. Verschelde, and A. Zhao. Newton's method with deflation for isolated singularities of polynomial systems. *Theoretical Computer Science*, 359(1):111–122, 2006.

[116] A. Leykin, J. Verschelde, and A. Zhao. Higher-order deflation for polynomial systems with isolated singular solutions. In *Algorithms in algebraic geometry*, pages 79–97. Springer, 2008.

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

[118] F. S. Macaulay. Some properties of enumeration in the theory of modular systems. *Proceedings of the London Mathematical Society*, 2(1):531–555, 1927.

[119] F. S. Macaulay. *The algebraic theory of modular systems*, volume 19. Cambridge University Press, 1994.

[120] A. Mantzaflaris and B. Mourrain. Deflation and certified isolation of singular zeros of polynomial systems. In *ISSAC'11*, page 249–256. ACM Press, 2011.

[121] A. Mantzaflaris, B. Mourrain, and A. Szanto. A certified iterative method for isolated singular roots. *Journal of Symbolic Computation*, 115:223–247, 2023.

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

[123] M. G. Marinari and T. Mora. Cerlienco-Mureddu correspondence and Lazard structural theorem. *Investigación Operacional*, 27(1):75–98, 2013.

[124] E. Mehrabi and É. Schost. A softly optimal Monte Carlo algorithm for solving bivariate polynomial systems over the integers. *Journal of Complexity*, 34:78–128, 2016.

[125] H. M. Möller and F. Mora. Upper and lower bounds for the degree of Gröbner bases. In *International Symposium on Symbolic and Algebraic Manipulation*, pages 172–183. Springer, 1984.

[126] B. Mourrain. Computing the isolated roots by matrix methods. *Journal of Symbolic Computation*, 26(6):715–738, 1998.

[127] S. Naldi and V. Neiger. A divide-and-conquer algorithm for computing gröbner bases of syzygies in finite dimension. In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, pages 380–387, 2020.

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

[129] V. Neiger, B. Salvy, É. Schost, and G. Villard. Faster modular composition, 2021.

[130] V. Neiger and É. Schost. Computing syzygies in finite dimension using fast linear algebra. *Journal of Complexity*, 60, 2020.

[131] R Notari and M. L. Spreafico. A stratification of Hilbert schemes by initial ideals and applications. *Manuscripta Mathematica*, 101:429–448, 2000.

[132] T. Ojika, S. Watanabe, and T. Mitsui. Deflation algorithm for the multiple roots of a system of nonlinear equations. *Journal of Mathematical Analysis and Applications*, 96(2):463–479, 1983.

[133] M. Olsson. *Algebraic spaces and stacks*, volume 62. American Mathematical Soc., 2016.

[134] K. Pardue. *Nonstandard Borel-fixed ideals*. Brandeis University, 1994.

[135] F. Pauer. On lucky ideals for Gröbner basis computations. *Journal of Symbolic Computation*, 14(5):471–482, 1992.

[136] S. R. Pope and A. Szanto. Nearest multivariate system with given root multiplicities. *Journal of Symbolic Computation*, 44(6):606–625, 2009.

[137] A. Ranum. The general term of a recurring series. *Bulletin of the American Mathematical Society*, 17(9):457–461, 1911.

[138] E. Riehl. *Category theory in context*. Courier Dover Publications, 2017.

[139] L. Robbiano. On border basis and Gröbner basis schemes. *Collectanea Mathematica*, 60:11–25, 2009.

[140] F. Rouillier. Solving zero-dimensional systems through the Rational Univariate Representation. *Applicable Algebra in Engineering, Communication and Computing*, 9(5):433–461, 1999.

[141] F. Rouillier. On solving systems of bivariate polynomials. In *ICMS*, volume 6327 of *Lecture Notes in Computer Science*, pages 100–104. Springer, 2010.

[142] É. Schost. Computing parametric geometric resolutions. *Applicable Algebra in Engineering, Communication and Computing*, 13(5):349–393, 2003.

[143] É. Schost and C. St-Pierre. Newton iteration for lexicographic gröbner bases in two variables, 2023. preprint on arXiv:2302.03766.

[144] M. Sherman. On an extension of galligo's theorem concerning the Borel-fixed points on the Hilbert scheme. *Journal of Algebra*, 318(1):47–67, 2007.

[145] V. Shoup. NTL: A library for doing number theory. `http://www.shoup.net`.

[146] V. Shoup. Fast construction of irreducible polynomials over finite fields. *Journal of Symbolic Computation*, 17(5):371–391, 1994.

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

[148] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo. *Force control.* Springer, 2009.

[149] A. Storjohann. Computation of Hermite and Smith normal forms of matrices. Master's thesis, University of Waterloo, 1994.

[150] A. Storjohann. *Algorithms for matrix canonical forms.* PhD thesis, ETH, Zürich, 2000.

[151] T. Tao. A trivial remark about schemes, 2012.

[152] C. D. Toth, J. O'Rourke, and J. E. Goodman. *Handbook of discrete and computational geometry.* CRC press, 2017.

[153] W. Trinks. On improving approximate results of Buchberger's algorithm by Newton's method. *SIGSAM Bull.*, 18(3):7–11, 1984.

[154] R. Vakil. The rising sea: Foundations of algebraic geometry notes, 2017.

[155] F. Winkler. A p-adic approach to the computation of Gröbner bases. *Journal of Symbolic Computation*, 6(2/3):287–304, 1988.

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

[157] X. Wu and L. Zhi. Determining singular solutions of polynomial systems via symbolic–numeric reduction to geometric involutive forms. *Journal of Symbolic Computation*, 47(3):227–238, 2012.

[158] N. Yamamoto. Regularization of solutions of nonlinear equations with singular Jacobian matrices. *J. Infor. Processing*, 7:16–21, 1984.

APPENDICES

Appendix A

Products

This section is a pure reminder of the definitions related to the exterior product. To learn more about those products please see [61, §A.2, A.3]

A.1 Tensor product

Let R be a ring, M, N be R-module and G be an abelian group with a group morphism $\varphi : M \times N \to G$. Then φ is R-**bilinear** if $\forall \lambda \in R$, $\forall m_i \in M$ and $n_i \in N$

- $\varphi(m_1\lambda, n_1) = \varphi(m_1, \lambda n_1)$
- $\varphi(m_1 + m_2, n_1) = \varphi(m_1, n_1) + \varphi(m_2, n_1)$
- $\varphi(m_1, n_1 + n_2) = \varphi(m_1, n_1) + \varphi(m_1, n_2)$

❧ **Definition 30.** *Let R be a ring and let M, N be R-module, the **tensor product** of M and N ($M \otimes N$) is F/U, where F the free abelian group on $M \times N$*

$$F = \bigoplus_{(m,n)\in M\times N} \mathbb{Z}e_{(m,n)}$$

and U is a $\mathbb{Z}$-submodule generarted by

- $e_{(m_1\lambda,n_1)} - e_{(m_1,\lambda n_1)}$
- $e_{(m_1+m_2,n_1)} - e_{(m_1,n_1)} - e_{(m_2,n_1)}$
- $e_{(m_1,n_1+n_2)} - e_{(m_1,n_1)} - e_{(m_1,n_2)}$

$\forall \lambda \in \mathbb{Z}$, $\forall m_i \in M$ *and* $n_i \in N$

In particular, the tensor product satisfies the universal property that if G is an abelian group with a bilinear group morphism $M \times N \to G$ then it factors through $M \otimes N$ in a unique manner

$$\begin{tikzcd}
M \times N \arrow[r] \arrow[rd] & M \otimes N \arrow[d, dashed] \\
 & G
\end{tikzcd}$$

A.2 Exterior product

The exterior product ($\wedge$) is defined using the tensor product.

❧ **Definition 31.** *Let R be a ring and let M be R-module, the* ***exterior product*** *of M ($M \wedge M$) is defined as $M \otimes M/I$, where $I = \bigcup_{m \in M} \langle m \otimes m \rangle$ is a two sided skew ideal.*

Remark A.2.1. The exterior product is anticommutative: $\forall a, b \in M$ then

$$0 \equiv (a+b) \wedge (a+b) \equiv (a \otimes a) + (a \otimes b) + (b \otimes a) + (b \otimes b) \equiv (a \otimes b) + (b \otimes a) \mod I.$$

Appendix B

Special Matrices

B.1 Hankel matrix

A thorough study of Hankel and Howell normal form matrices and their computation can be found in [150, 149].

❧ **Definition 32.** *An **Hankel matrix** is a square symmetric matrix $\boldsymbol{A}$ such that $\boldsymbol{A}[i,j] = \boldsymbol{A}[i-l, j+l]$ for all l such that $\boldsymbol{A}[i-l, j+l]$ is not out of bound [20][§11].*

B.2 Hermite normal form

Let $\mathbb{A}$ be a principal ideal domain (PID) and let $\mathbb{A}^{n\times m}$ be the set of $n \times m$ matrices over $\mathbb{A}$.

❧ **Definition 33.** *We say that $\boldsymbol{A} \in \mathbb{A}^{n\times m}$, then $\boldsymbol{A}[i,j]$ is an **upper left pivot** of $\boldsymbol{A}$ if*

- $\boldsymbol{A}[i,j] \neq 0$
- $\boldsymbol{A}[k,l] = 0$ *for* $k \geq i$ *and* $l \leq j$ *and* $(k,l) \neq (i,j)$

❧ **Definition 34.** *We say that $\boldsymbol{A} \in \mathbb{A}^{n\times m}$ is in **echelon reduced form** if $\boldsymbol{A}$ is in an upper triangular form, that is*

- *all non zero row $\boldsymbol{A}[i,-]$ in $\boldsymbol{A}$ have an upper left pivot $\boldsymbol{A}[i,j_i]$ for some j_i;*
- *if $\boldsymbol{A}[i,-]$ is a zero row, then $\boldsymbol{A}[l,-]$ is a zero row $\forall l > i$;*
- *$\boldsymbol{H}[l,j_i] \in C_i$ for all $l < i$ where C_i is a complete set of associates of $\mathbb{A}$.*

Note that the definition of an upper left pivot ensures $j_i > j_l$ for all $l > i$ which guarantees a triangular upper form.

❧ **Definition 35.** *Let $\mathbb{A}$ be a domaine and $\mathbb{K}$ its fraction field. Given $A \in \mathbb{K}^{n\times m}$ with entries in $\mathbb{A}$, then $\boldsymbol{H}$ is the* ***Hermite normal form*** *of the matrix $\boldsymbol{A}$ if*

- *$\boldsymbol{A}$ and $\boldsymbol{H}$ are left equivalent, i.e. $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{A}$ for $\boldsymbol{H} \in \mathbb{A}^{n\times m}$ for $\boldsymbol{U} \in \mathbb{K}^{n\times n}$ with $\det(\boldsymbol{U}) \in \mathbb{A}^{\times}$ (unimodular);*

- *if $\boldsymbol{A}$ is in the echelon reduced from;*

A Hermite normal of $\boldsymbol{A}$ always exists [1, §5, theorem 2.9] and is unique for a fix set of associates [1, §5 , theorem 2.13] . The Hermite normal form can be defined on rows or columns. The above definition and this document use columns, but all the definitions can be transposed to define the Hermite normal form on the rows.

B.3 Howell normal form

Howell matrix is analogous to Hermite matrices for *principal ideal ring* (PIR). Let $\mathbb{A}$ be a PIR and let $\mathbb{A}^{n\times m}$ be the set of $n \times m$ matrices over $\mathbb{A}$. Given $\boldsymbol{A} \in \mathbb{A}^{n\times m}$ with entries in $\mathbb{A}$, then $\boldsymbol{H}$ is the **Howell normal form** of the matrix $\boldsymbol{A}$ if

- $\boldsymbol{A}$ and $\boldsymbol{H}$ are left equivalent;

- if $\boldsymbol{H}$ is in the echelon reduced from;

- Let $\mathcal{R}$ be the row span of $\boldsymbol{A}$ then for $r \in \mathcal{R}$ if $r[i] = 0$ for all $i < j$ then r is in the $\mathbb{A}$-span of the j last rows of $\boldsymbol{H}$.

Again, the definition can be transposed to define the Howell normal form on the rows. But we adopted the column convention in this document.

Appendix C

Gröbner bases applications

We now review some other additional motivations to find Gröbner bases. There are many, so we sampled few common applications in algebra for this section and, in the next section, we show some motivation from geometry. Some extended lists of applications can be found in [45], [44], [61], and [58].

Remark C.0.1. We also review some applications of bases for orderings that are not lexicographic. This motivates the short discussion concerning different orderings.

C.1 Applications in algebra

Here are a few examples of Gröbner bases in algebra.

❧ **Example 3.1.1: Membership problem**

Proposition C.1.1. *[61, §15.1.1] Let $f \in \mathbb{K}[x_1, \ldots, x_n]$ and I be some ideal over $\mathbb{K}[x_1, \ldots, x_n]$ generated by the Gröbner basis $g_1, \ldots, g_t$. Then $f \in I$ if and only if the remainder r upon the division by $g_1, \ldots, g_t$ is 0.*

The proof is simple and worth summarizing since we discuss division with a Gröbner basis in Chapter 4

Proof. Since $g_1, \ldots, g_t$ form a Gröbner basis of I, we have $\mathbf{In}(I) = \langle in(g_1), \ldots, in(g_t)\rangle$. Reusing the symbols from the division algorithm, we have $r = f - (q_1 g_1 + \cdots + q_t g_t)$ so r is in I if and only if f is since the g_i are in I, and that for any step. This proves the reverse direction. For the forward direction, we assume $f \in I$, then by the statement on r from above, the first stopping condition never occurs. It follows that r must become 0 at some points since the algorithm always terminates.[61, §15.1.1] □

Gröbner bases and Buchberger's algorithm are also helpful to find syzygies and free resolutions of modules.

❧ Example 3.1.2: §Syzygies

Buchberger's algorithm also presents a second utility: the S-polynomials, which remainders vanish by Buchberger's criterion, give the syzygies on a Gröbner basis (by *Schreyer modifications*).

Proposition C.1.2 (Schreyer)**.** *[61, theorem 15.10] Let $\mathcal{M}$ be a monomial submodule of R^t with a Gröbner basis $\mathcal{G} = (g_1, \ldots, g_t)$, for any monomial ordering. Then the S-polynomials of the g_i's are a Gröbner basis of the first syzygy of $g_1, \ldots, g_t$.*

By extension, Gröbner bases may also be used to find a finite free resolution for a submodule of R^d, which always exists by Hilbert's syzygy theorem.

❧ **Definition 36.** *Let $\mathcal{M}$ be a module. A finite free resolution is an exact sequence:*

$$0 \to F_t \to \cdots \to F_1 \to \mathcal{M} \to 0.$$

where the F_i's are free modules.

Theorem C.1.1 (§ Hilbert's syzygy theorem)**.** *[108, §XXI 4, theorem 4.15] Let $\mathcal{M}$ be a finely generated $\mathbb{K}[x_1, \ldots x_n]$-module, then there exists a finite free resolution of $\mathcal{M}$. That is we can find a finite exact sequence*

$$0 \to F_m \to F_{m-1} \to \cdots \to F_1 \to F_0 \to \mathcal{M}$$

where $m \leq n$ and the F_i's are free.

❧ Example 3.1.3: Free resolutions of modules

Hereinbefore, we showed that the S-polynomials of a Gröbner basis $\mathcal{G}$ of a module $\mathcal{M}$ can be used to get a Gröbner basis of the first syzygies of $\mathcal{G}$: $syz(\mathcal{G})$. Repeating the process on the Gröbner basis of $syz(\mathcal{G})$, we find a second syzygy $syz(syz(\mathcal{G}))$ and so on. Hence, a Gröbner basis can be used to build free resolutions of modules

$$0 \to R^{t_m} \ldots \xrightarrow{\varphi_3} R^{t_1} \xrightarrow{\varphi_2} R^{t_0} \xrightarrow{\varphi_1} \mathcal{M} \to 0$$

where $\ker(\varphi_i)$ is the i^{th} syzygy of $g_1, \ldots, g_t$. The procedure always reaches a free module of $m \leq n$ by Hilbert's syzygy theorem.

❧ **Example 3.1.4: Homogenization**

With the graded ordering, an ideal can easily be homogenized. If $I \subseteq R$ is an ideal of a polynomial ring R with a gröbner basis $\mathcal{G}$, then the homogenization of I is

$$I^h = \langle \{g^h \mid g \in \mathcal{G}\} \rangle$$

where g^h is the homogenized polynomial [45, §8.4, theorem 4].

C.2 Connections to geometry

There exist many applications in geometry; we just review some here, see [61, §15][45, 44] for more examples. Henceforward, we fix $\mathbb{K} = \overline{\mathbb{K}}$ and we are only interested in Gröbner bases of ideals over polynomial rings $\mathbb{K}[x_1, \ldots, x_n]$. The Gröbner bases are now assumed to be reduced and minimal with $in(g_i) \succ in(g_{i+1})$.

❧ **Example 3.2.1: Elimination**

Let $g_1, \ldots, g_t$ be a lexicographic $(x_i \succ x_{i+1})$ Gröbner basis of an $I \subset \mathbb{K}[x_1, \ldots, x_n]$ then $I_l = I \cap \mathbb{K}[x_1, \ldots, x_{l-1}]$ is generated by the g_i such that $x_l \succ in(g_i)$.

Proof. See [45, §3 theorem 2] □

In geometry, this results in a *projection* of $V(I)$ in $V(\langle x_l, \ldots, x_n \rangle)$.

C.2.1 A Word on Dimension

Although our context is restricted to zero-dimensional ideals, it is worth mentioning a few connections between the Gröbner basis and dimension theory.

❧ **Definition 37** (Krull dimension)**.** *The Krull dimension of a ring $\mathbb{A}$ ($\dim \mathbb{A}$) is the length of the longest chain (supremum) of prime ideals in $\mathbb{A}$. For $\mathcal{M}$ a R module the Krull dimension of $\mathcal{M}$ is the Krull dimension of $R/Ann(\mathcal{M})$ as a ring where $Ann(\mathcal{M}) := \{r \in R \mid rm = 0, \forall m \in \mathcal{M}\}$ is the annihilato r of $\mathcal{M}$ [61, §8].*

In geometry the Krull dimension of $\mathbb{K}[x_1, \dots, x_n]/I$ is

- the transcendence degree $\mathbb{K}[x_1, \dots, x_n]/I$ over $\mathbb{K}$ [61, §8];
- the dimension (dim) of a variety $V(I)$ if $\mathbb{K} = \bar{\mathbb{K}}$ [45, §9.3, Theorem 8].

We abuse notation and say I is dimension m when $\mathbb{K}[x_1, \dots, x_n]/I$ has dimension m.

❧ **Example 3.2.2**

$\dim(\mathbb{K}) = 0$ and $\dim(\mathbb{K}[x_1, \dots, x_n]) = n$

❧ **Example 3.2.3: Zero dimensional ideal**

If $I \subsetneq R$ is zero-dimensional, then $\operatorname{Spec} R/I$ only contains maximal ideals and, by Nulstellensatz, $|V(I)| \leq \infty$. For example, all varieties in Example 1.1.1 are zero-dimensional.

Connections with the dimension theory can be viewed using the Hilbert function. The Hilbert function is a nice asset to describe the dimension of a graded vector space as they compress the information of the dimension of the direct summands of the grading.

❧ **Definition 38.** *Let $R = \mathbb{K}[x_1, \dots, x_n]$ and $I \subseteq R$ be and ideal. We define $R_{\leq u}$ to be the set of monomials of degree smaller than u and $I_{\leq u} = I \cap R_{\leq u}$. Then we can define a function subject to the constaint $HF_{R/I}(u) = dim_{\mathbb{K}}(R_{\leq u}/I_{\leq u})$, for u large. Such function is called the **Hilbert function** of R/I.*

❧ **Example 3.2.4: Hilbert function of univaritate polynomial ring**

$\mathbb{K}[x] \cong \mathbb{K} \oplus \mathbb{K}x \oplus \mathbb{K}x^2 \oplus \dots$, hence $HF_{\mathbb{K}[x]}(u) = 1 + u + u^2 + \cdots = \frac{1}{1-u}$

❧ Example 3.2.5

Let $deg(x_i) = a_i$. Then $HF_{\mathbb{K}[x_1,\ldots,x_n]}(u) = (1 + u^{a_1} + u^{2a_1} + \ldots) \ldots (1 + u^{a_n} + u^{2a_n} + \ldots) = \frac{1}{\prod_{i=1}^{n}(1-u)^{a_i}}$.

The Hilbert function can be found using Gröbner bases with a graded ordering.

Proposition C.2.1. *Let* $\mathbb{K} = \bar{\mathbb{K}}$ *and let* $R = \mathbb{K}[x_1, \ldots, x_n]$ *with* $I \subseteq R$ *an ideal For a graded ordering,* $\dim V(I) = HP_{R/I}(u) = HP_{R/\mathbf{In}(I)}(u)$ *for* u *large [45, §9.3, Theorem 8].*

Hence when finding the dimension of a variety of an ideal I, one need only consider monomial ideals $\mathbf{In}(I)$, which reduces the problem to a combinatorics problem.

❧ Example 3.2.6: Transverse Intersection

Let $I = \langle x^2 + 2x - y + 1, (y-2)^2 \rangle \subseteq \mathbb{K}[x, y]$ be an ideal and its Gröbner basis for $\succ_{grevlex}$. By Proposition C.2.1, $HP_{\mathbb{K}[x,y]/I}(u) = HP_{\mathbb{K}[x,y]/\langle x^2,y^2 \rangle}(u)$. It follows that

$$dim(\mathbb{K}[x, y]/I) = dim(\mathbb{K}[x, y]/\langle x^2, y^2 \rangle) = 0.$$

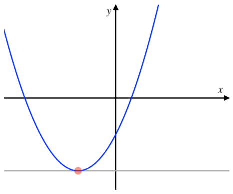
❧ Example 3.2.7: §Twisted Cubic (complete intersection in $\mathbb{A}^3$)

Consider the ideal $I = \langle y^2 - xz, yz - x, z^2 - y \rangle \subseteq \mathbb{K}[x, y, z]$, an ideal and its Gröbner basis for $\succ_{grevlex}$, *i.e.*, the twisted cubic in $\mathbb{A}^3$. By Proposition C.2.1, we infer that since $V(I)$ has dimension 1 since

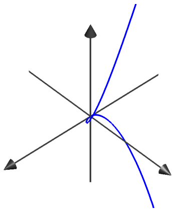
$V(\langle z^2, zy, y^2 \rangle) = V(\langle z^2 \rangle) \cap V(\langle zy \rangle) \cap V(\langle y^2 \rangle) = V(\langle z^2 \rangle) \cap V(\langle y^2 \rangle) =$ the x-axis.

❧ Example 3.2.8: §Twisted Cubic (non complete intersection in $\mathbb{P}^3$)

Consider the ideal $I = \langle y^2 - xz, yz - xw, z^2 - yw \rangle \subseteq k[x, y, z, w]$ (Gröbner basis in $\succ_{grevlex}$) where I is the twisted cubic in $\mathbb{P}^3$. By Proposition C.2.1, $dim(y^2 - xz, yz - xw, z^2 - yw) = dim(z^2, zy, y^2)$ and $V(z^2, zy, y^2) = V(z) \cap V(y)$. Since $V(z) \cap V(y)$ is a line in $\mathbb{P}^3$, it follows that $V(I)$ is also a line in $\mathbb{P}^3$.



$V(x^2 + 2x - y + 1, (y-2)^2)$



$V(y^2 - xz, yz - xw, z^2 - yw)$

So Gröbner bases can greatly simplify the calculation of the dimensions. More examples can be found in [45, §9]. Since this thesis only focuses on zero-dimensional ideals, let us review a last application of the lexicographic ordering.

Theorem C.2.1 (Finiteness Theorem)**.** *Let $I \subseteq R$ and ideal, then the following are equivalent:*

(a) R/I is a finite-dimensional $\mathbb{K}$-vector space;

(b) the Krull dimension of $\mathbb{K}[x_1, \ldots, x_n]/I$ is 0;

(c) for all $i = 1, \ldots, n$ there exists $a_i \in \mathbb{N}^+$ such that $x_i^{a_i}$ is the initial term of a generator in the lexicographic Gröbner basis of I.

A proof of Theorem C.2.1 can be found in [44, §5.3, Theorem 6]. In Section 2.1.7, the size of the moduli space of a stratum, and therefore the number of parameters, is related to the number of monomials under the staircase. The above theorem further enforces that the number of monomials is finite.