Regularisation of Gröbner basis computations for weighted and determinantal systems, and application to medical imagery

Thibaut Verron

UPMC
SORBONNE UNIVERSITÉS

THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Informatique

École Doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Thibaut Verron

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Régularisation du calcul de bases de Gröbner pour des systèmes avec poids et déterminantiels, et application en imagerie médicale

Thèse dirigée par Jean-Charles Faugère et Mohab Safey El Din

soutenance prévue le **lundi 26 septembre 2016**

après avis des **rapporteurs**:

M. Laurent Busé	Chargé de recherche, Inria, HDR
M. Bruno Salvy	Directeur de recherche, Inria

devant le **jury** composé de :

M. Jean-Charles Faugère	Directeur de recherche, Inria
M. Mohab Safey El Din	Professeur, Université Pierre et Marie Curie
M. Laurent Busé	Chargé de recherche, Inria, HDR
M. Bruno Salvy	Directeur de recherche, Inria
M. Bernard Bonnard	Professeur, Université de Bourgogne
M. Stef Graillat	Professeur, Université Pierre et Marie Curie

Contents

Introduction

1. Motivation

1.1. Polynomial system solving and applications

The problem of solving polynomial systems is very important nowadays, because such systems arise in a wide range of applications. Indeed, polynomials are a basis for models in many areas of research.

We give three examples of general applications, both practical and theoretical, of polynomial system solving that we will encounter in this thesis.

Cryptography Algebraic cryptanalysis models cryptographical systems using polynomial systems, and reduces the problem of breaking the code to solving these systems. We shall encounter an example related to the Discrete Logarithm Problem (DLP) on elliptic curves.

Polynomial inversion The problem is, given some polynomials $f_1, \ldots, f_m$, to identify the relations between the f_i's. For example, the only nontrivial relation between $f_1 = XY$, $f_2 = X^2$ and $f_3 = Y^2$ is $f_1^2 = f_2 f_3$. It can be used for example to obtain implicit equations describing a parametric object.

Real roots classification This is also related to parametric equations: the input is a polynomial system $F = (f_1(\mathbf{X}, \mathbf{T}), \ldots, f_m(\mathbf{X}, \mathbf{T}))$ with coefficients in $\mathbb{R}$, indeterminates $\mathbf{X} = (X_1, \ldots, X_n)$ and parameters $\mathbf{T} = (T_1, \ldots, T_s)$ such that for any $\mathbf{t} \in \mathbb{R}^s$, the equations $(f_1(\mathbf{X}, \mathbf{t}), \ldots, f_m(\mathbf{X}, \mathbf{t}))$ with unknowns $\mathbf{X} = (X_1, \ldots, X_n)$ have a finite number of real solutions C_t. The goal is to find a dense covering of $\mathbb{R}^s$ with open subsets over which C_t is constant.

These examples illustrate the fact that the meaning of *solving* a polynomial system depends on the application. The simplest question one might ask given a polynomial system is "*does it have a solution?*". This problem is NP-complete on finite fields. Solving a system usually means computing some more information about the solution set. The exact form of the question depends on the application, and systems are usually split in two categories: they can be *zero-dimensional* or *positive-dimensional*.

Zero-dimensional systems are systems with a finite number of solutions, in an algebraic closure of the coefficient field. For example, systems arising in applications to cryptography are usually zero-dimensional: there are enough constraints to ensure that there is only one solution to the problem. In this case, "solving" usually means either listing all the solutions of the system, giving a parametrisation of all solutions of the system, or simply giving one solution of the system.

On the other hand, positive-dimensional systems are systems with infinitely many solutions, again in some algebraic closure of the coefficient field. There, "solving" potentially takes on many

different meanings, depending on the system and the application: finding a parametrisation of the solutions, computing the equations of the projection of the solutions on a subspace (or in other words, eliminate variables from the system), computing the dimension of the solution set (that is identifying a subset of unknowns which are independent), or, when applicable given the coefficient field, finding topological information on the set of solutions (for example one point per connected component of the set or its complementary).

For example, the polynomial inversion problem is an elimination problem: if the input polynomials are $(f_1, \ldots, f_m) \in \mathbb{K}[X_1, \ldots, X_n]$, one can construct the ideal $I = \langle F_1 - f_1, \ldots, F_m - f_m \rangle$ with extra variables $F_1, \ldots, F_m$, and the wanted relations shall correspond to polynomials in $I \cap \mathbb{K}[F_1, \ldots, F_m]$, thus eliminating $X_1, \ldots, X_n$.

The real roots classification problem also involves an elimination problem: we need to find the set of values of the parameters where the number of real solutions in $\mathbf{X}$ changes. Let $\pi : \mathbb{R}^n \times \mathbb{R}^t \longrightarrow \mathbb{R}^s$ be the projection onto the parameters $T_1, \ldots, T_s$, the wanted subdivision of the parameter space is contained in the projection of the singular locus of the zeroes of F and of the critical points of π restricted to this variety. It is enough to compute one polynomial whose zeroes cover this set, and this too can be done using elimination. In order to complete the resolution of the problem, we then need to compute one point per connected component of the complementary of this set, and for each point, count the solutions in $\mathbf{X}$.

Many tools have been developed for solving polynomial systems, including numerical methods (Newton's method…) or symbolic-numeric methods (homotopy continuation [SW05]…), computing solutions with arbitrary precision, as well as symbolic methods from computer algebra (triangular sets [ALM99], multivariate resultants [CD05], geometric resolution [GLS01]…) computing exact solutions.

In this thesis, we mainly focus on Gröbner bases [Buc76]. They can be used to solve systems with a finite number of solutions as well as eliminate variables, thus covering most needs for polynomial system solving in applications. Furthermore, they can solve a polynomial system regardless of its field of definition.

We shall also make use of regular chains [LMX05], as an alternative tool for computing polynomial elimination, and Cylindric Algebraic Decompositions [Col75] in order to obtain topological information about sets of real solutions of a system of polynomial equations and inequations.

1.2. Regularity of Gröbner basis computations

Gröbner bases are computed by successive reductions of the polynomials defining the system. This can be done in an algebraic way, similar to euclidean reductions in the univariate case, or using linear algebra on the coefficients of the polynomials. The historical algorithm by Buchberger [Buc76] uses the former method, while modern algorithms such as F_4 [Fau99] or F_5 [Fau02] use linear algebra to group reductions together. This makes the choice of a *reduction strategy* important, to condition how, at any given step, to pick polynomials for the reduction.

The default strategy is the *normal strategy* for homogeneous systems (or its inhomogeneous variant, the *sugar strategy*), which picks at each step the polynomials leading to the smallest degree reductions. For this strategy, the degree of the reductions can be seen as an indicator of progress, and we can characterize regular behaviors for the algorithms: a behavior is regular if

it does show any *degree fall*, that is a step where the degree of the reductions is not (strictly) increasing. See for example the curves on page 10: the red curve is an example of a regular behavior, whereas the blue curve is an example of an irregular behavior.

We shall see that regular behaviors make it possible to analyze the complexity of the algorithms: indeed, if at each step the degree is strictly increasing, we may bound the number of steps by bounding the maximal degree, and then obtain complexity bounds by bounding the cost of each step.

This can be done under some *generic properties*, that is algebraic properties which are satisfied by *almost all* systems. An example of a generic property relevant for Gröbner basis computations is zero-dimensional regular sequences, that is sequences of n polynomials in n variables having only a finite number of solutions (in an algebraic closure of the coefficient field). For homogeneous zero-dimensional regular sequences, or sequences such that the highest degree homogeneous components form a zero-dimensional regular sequence, algorithm F_5 has a regular behavior, and algorithm F_4 has as few degree falls as possible.

1.3. What to do when systems are not generic?

Problems appear when systems fail to satisfy genericity properties: complexity bounds relying on the regular behavior of the algorithms are no longer valid, and the algorithms waste time computing useless reductions.

For example, the authors of [Fau+13], when working with the DLP on Edwards elliptic curves, were confronted with the following system of 5 equations in 5 unknowns, over a finite field:

$$\begin{aligned} 0 = &\begin{bmatrix}7871\\18574\\14294\\32775\\20289\end{bmatrix} e_5^{16} + \begin{bmatrix}53362\\50900\\36407\\58813\\20802\end{bmatrix} \tilde{e}_1^{8} + \begin{bmatrix}26257\\128\\3037\\38424\\41456\end{bmatrix} \tilde{e}_1^{7}\tilde{e}_2 + \begin{bmatrix}25203\\23117\\28918\\29298\\56353\end{bmatrix} \tilde{e}_1^{6}\tilde{e}_2^{2} + \begin{bmatrix}19817\\29737\\52187\\36574\\46683\end{bmatrix} \tilde{e}_1^{5}\tilde{e}_2^{3} + \begin{bmatrix}9843\\3752\\27006\\64195\\63059\end{bmatrix} \tilde{e}_1^{4}\tilde{e}_2^{4} \\ &+ \begin{bmatrix}11204\\25459\\58263\\17964\\57146\end{bmatrix} \tilde{e}_1^{3}\tilde{e}_2^{5} + \begin{bmatrix}46217\\5478\\45631\\13171\\42548\end{bmatrix} \tilde{e}_1^{2}\tilde{e}_2^{6} + \begin{bmatrix}63811\\50777\\48809\\1858\\55751\end{bmatrix} \tilde{e}_1\tilde{e}_2^{7} + 2070 \text{ smaller monomials.} \end{aligned}$$

This system is not a zero-dimensional regular sequence: the highest degree component of all 5 polynomials is a monomial in e_5^{16}, so any point with $e_5 = 0$ is a root of these components. And indeed, the behavior of Gröbner basis algorithms using the normal strategy for this system is irregular: see on Figure 1 the blue plot, representing the progress of algorithm F_5, step by step. Each step at which the plot stalls or reverses course is an irregularity, making it harder to bound the complexity.

On the other hand, assume that we now assign degree 2 to variables $\tilde{e}_i$, $i \in \{1, \ldots, 4\}$ (thus excluding e_5). In other words, we are using the *weighted degree* for the system of weights $(2, 2, 2, 2, 1)$. Then all monomials printed above have weighted-degree 16, together with 486 more monomials. The highest weighted homogeneous components of polynomials of the system

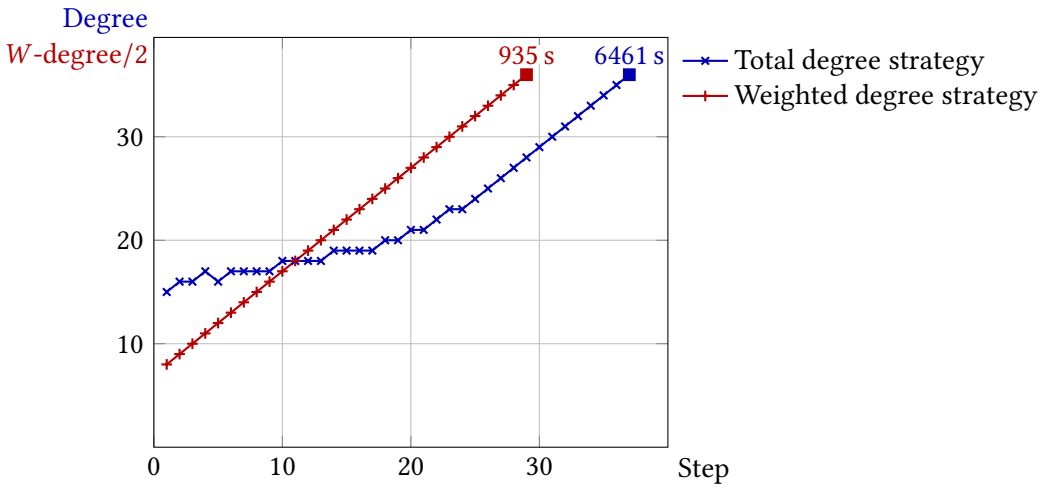

Figure 1: Step-by-step progress of Algorithm F_5 on a DLP system

are now much larger than just one monomial, and they actually form a zero-dimensional regular sequence. By evaluating the weighted degree of polynomials for the reduction strategy, instead of the total degree, we recover a regular behavior. The behavior of algorithm F_5 using the weighted degree was plotted in red on Figure 1, and indeed the algorithm progresses steadily towards its final weighted degree.

Another example is given by determinantal ideals: a *determinantal system* is a system defined as the set of all r-minors of a matrix with polynomial entries. It encodes the fact that the matrix has rank at most $r-1$. These systems appear in a lot of applications, for example in cryptography (the MinRank problem [FLP08]) or in convex optimization (semi-definite programming [Ott+15]). Computing critical points of projections restricted to a determinantal variety is a frequent problem, which has been thoroughly studied ([BV88]...).

For example, consider a $k \times k$ matrix M with coefficients in $\mathbb{K}[X_1, \ldots, X_n]$, and let V be the variety defined by $\det(M)$ (that is the $k-1$'th determinantal variety associated with M). Assume that we want to compute the critical points of the projection onto the first variable X_1, restricted to V.

A natural way of approaching this problem is to use the Jacobian criterion on the determinantal system: compute the partial derivatives $\frac{\partial \det(M)}{\partial X_i}$, $i \in \{2, \ldots, n\}$, saturate by $\frac{\partial \det(M)}{\partial X_1}$, and compute a Gröbner basis of this system. In the same way as before, we plot the behavior of this strategy in blue in Figure 2, for $k = 5$, $n = 7$ and $\mathbb{K} = \mathbb{F}_{65521}$. Again, it appears that there are some irregularities.

However, it is possible to change the modelization, or in other words find another set of equations describing the same geometrical object. The idea is that the matrix M has rank at most $r-1$ if and only if its kernel has dimension at least $k-r+1$. This can be modelled by adding $k(k-r+1)$ variables to the polynomial algebra, form a $k \times (k-r+1)$-matrix N with these variables as entries, and use the entries of $M \cdot N$ as the polynomial systems. We want N

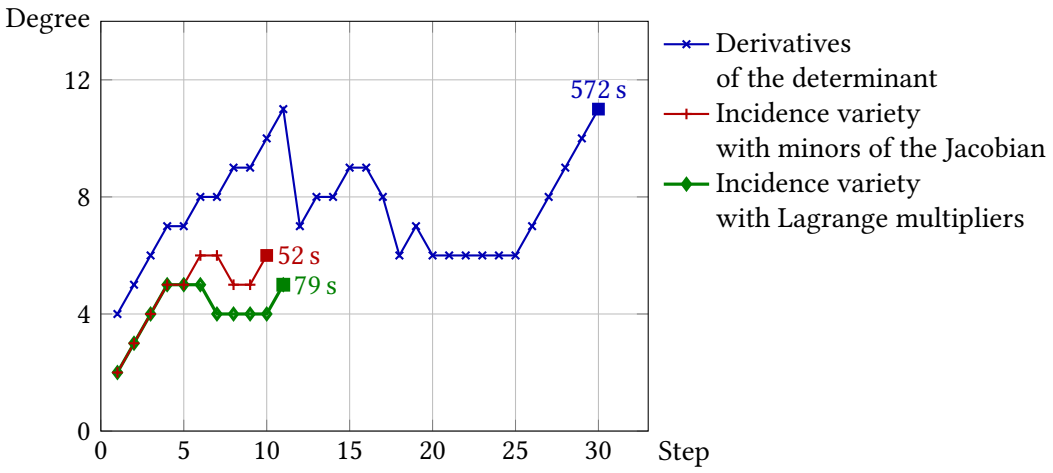

Figure 2: Step-by-step progress of Algorithm F_4 on the singularities of a determinantal system

to have full rank $(k - r + 1)$, this can be described locally with $(k - r + 1)^2$ additional equations.

The resulting system F_I defines what is called an *incidence variety*. It models the same geometrical object, and we can compute the critical points of the projection restricted to this variety using the Jacobian criterion: critical points of the projection onto X_1 are points where the truncated Jacobian matrix $\mathrm{Jac}_{X_2,\ldots,X_n,Y_1,\ldots,Y_{k(k-r+1)}}(F_I)$ has rank less than the codimension of the variety.

This is a determinantal condition too, and we can model it with minors of the Jacobian, or again form an incidence variety using the Jacobian. This last approach is known in optimization as the introduction of *Lagrange multipliers*. In this case, both these approaches make the computations more regular, and faster by a factor of almost 10. On Figure 2, we plotted in red (*resp.* green) the behavior of the computations on the minors of the Jacobian of the incidence system (*resp.* the system obtained by multiplying the Jacobian with Lagrange multipliers), with the same input as before.

Both weighted-homogeneous and determinantal systems are examples of the more general problem of solving *structured systems*: systems arising in applications frequently exhibit some structure that they do not share with arbitrary systems. This is both a problem and an advantage: on the one hand, this makes the systems non-generic, and usually the algorithms need to be adapted to recover a regular behavior for these systems; and on the other hand, once dedicated algorithms exist, the additional structure frequently makes it easier to solve generic structured systems than arbitrary generic systems.

In general, working with a new structure raises two deeply interleaved problems:

- identify properties of generic polynomials with the structure, which imply a regular behavior of dedicated algorithms;
- and design dedicated algorithms taking the structure into account and having a regular behavior for generic systems.

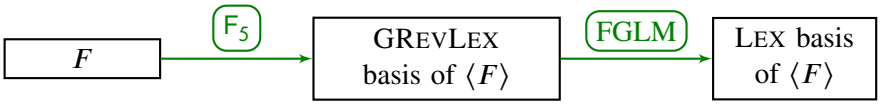

Figure 3: Algorithmic strategy for zero-dimensional systems

2. State of the art

2.1. Computational strategy for Gröbner bases

We now give more details about how Gröbner bases can be used for solving polynomial systems, and why a reduction strategy taking into account the degree of the polynomials is useful in this workflow.

The definition of a Gröbner basis depends on the choice of a *monomial ordering*, in order to give a formal meaning to the *reduction* of polynomials. Gröbner bases for different monomial orderings will yield different kinds of information, and be more or less easy to compute. In terms of actually solving the system, two monomial orderings are particularly useful:

- elimination orderings, that is orderings such that any monomial divisible by $X_1, \ldots, X_i$ is greater than any monomial not divisible by $X_1, \ldots, X_i$, for some i: using such an ordering, one may eliminate the variables $X_1, \ldots, X_i$ from the ideal;
- the lexicographical ordering: up to a random change of coordinates, the lexicographical Gröbner basis of a zero-dimensional system is a parameterized representation of all solutions: X_n is given as zero of a univariate polynomial, and all other unknowns are expressed as polynomials in X_n.

On the other hand, computing bases for elimination orderings or the lexicographical ordering is usually more difficult than for other orderings such as the GRevLex ordering. A common strategy for solving a polynomial system is thus to first compute a GRevLex basis using algorithm F_4 or F_5, and then perform a change of ordering to compute an elimination or lexicographical basis from the GRevLex basis. This change of ordering can be done using another run of algorithm F_4 or F_5, or using a dedicated algorithm such as FGLM [Fau+93] in the zero-dimensional case (Figure 3), or the Gröbner walk [CKM97] in the general case.

This makes computing a Gröbner basis for the GRevLex ordering an important step in most cases. This order orders the monomials according to their degree first, which is why algorithms usually use the normal strategy by default when computing a GRevLex basis: it is a way to consider the smallest monomials first.

2.2. Complexity studies for homogeneous systems

Homogeneous systems are maybe the most widely studied structure for Gröbner basis computations: indeed, from the very start, strategies such as the normal strategy or the sugar strategy were designed in order to take into account the degree of homogeneous polynomials, or properties of their highest degree components when the systems were inhomogeneous.

It has been observed that the normal strategy frequently gives good results in practice. This does not necessarily mean that the algorithms are lightning-fast, but monitoring their progress does not show any "irregular" behavior of the kind of our earlier examples: the degree of the polynomials considered grows steadily from start to end.

The study of this behavior led to algebraic characterizations of systems exhibiting this regular behavior. It turned out that these properties are *generic*, which is a formal way of expressing that "most" systems will satisfy them, and explains why these strategies appeared efficient in practice.

An example of generic property ensuring regularity in Gröbner basis algorithms is *regular sequences*. It has several equivalent definitions, but for now we will restrict to systems with n equations and n unknowns (so-called *square* systems). A square system is a regular sequence if and only if its set of solutions is finite.

This property ensures that algorithm F_5 has a regular behavior on homogeneous systems. For inhomogeneous systems, the "good" property is *regularity in the affine sense*, and it requires that the highest degree components of the system form a regular sequence.

Under these genericity hypotheses, the regular behavior of the algorithms makes it possible to obtain complexity bounds. For example, here are complexity bounds for algorithms F_5 and FGLM for generic square systems.

Theorem 1. *Let $F = (f_1, \ldots, f_n) \subset \mathbb{K}[X_1, \ldots, X_n]$ be a homogeneous system with respective degree $d_1, \ldots, d_n$. Assume that the system is regular in the affine sense (and thus zero-dimensional).*

Then algorithm F_5 computes a GRevLex Gröbner basis of $\langle F \rangle$ in time

$$O\left(d_{\text{reg}}\binom{n + d_{\text{reg}} - 1}{d_{\text{reg}}}^3\right),$$

where d_{reg} is the degree of regularity *of the system, that is the largest degree we need to reach in a run of F_5. It is bounded by Macaulay's bound [BFS14, Cor. 13]:*

$$d_{\text{reg}} \leq \sum_{i=1}^{n}(d_i - 1) + 1.$$

Algorithm FGLM computes a lexicographical basis of $\langle F \rangle$ in time

$$O\left(n \deg(I)^3\right),$$

where $\deg(I)$ *is the* degree *of the ideal, that is the number of solutions. It is given by Bézout's bound [Laz83, Prop. 1]:*

$$\deg(I) = \prod_{i=1}^{n} d_i.$$

2.3. Other structures

The problem of computing Gröbner bases for structured polynomial systems is not new, and a lot of new structures have been successfully exploited from this point of view in the past few

years: examples include multi-homogeneous systems [FSS11], systems invariant under a group action [FS12], determinantal systems [FSS13]...

In particular, this last example will prove useful in this thesis. Recall that given a $k \times k$-matrix M with polynomial entries, and an integer $r \in \{1, \ldots, k\}$, the *determinantal system* D_r is the system given by all r-minors of M. Its zeroes, the so-called *determinantal variety*, are points at which the matrix M has rank at most $r - 1$.

The structure of these ideals has been thoroughly studied, both from an algebraic standpoint [BV88] and from a computational standpoint [FSS13]. In particular, it is known that generically, the singular locus of D_r is the determinantal variety D_{r-1}.

As for computations, the idea of modelling the determinantal variety as the projection of an incidence variety is far from new: for example, the classic method of using Lagrange multipliers in optimization relies on modelling rank defects of the Jacobian matrix of a system as an incidence variety, and this technique was central in the work of [Nal15] on real optimization for determinantal varieties. This strategy has been analyzed from a computational point of view in the zero-dimensional case [FSS13].

3. Contributions

In this thesis, we worked on two classes of structured polynomial systems: weighted homogeneous systems, and determinantal systems for a root classification problem.

Weighted homogeneity is the structure used in the DLP example on page 9. It generalizes homogeneity, by giving more choice over how the degree is computed. This structure appears naturally in a wide range of applications, for example whenever one applies a polynomial change of variables to a homogeneous system. Algorithmic strategies for weighted homogeneous systems were known, but there was no complexity analysis. Furthermore, the FGLM step of the classic strategy became a bottleneck in the resolution. We identified genericity properties which allowed us to obtain complexity bounds for the algorithms, and described a strategy allowing algorithm FGLM to take advantage of the structure. Under genericity hypotheses, we proved that the algorithms have a completely regular behavior in this strategy.

On the other hand, for determinantal systems, we started with a specific problem from an application to contrast optimization, in medical imagery. While solving this problem, we identified that the underlying problem was the more general problem of real root classification, specialized to singularities of determinantal varieties. However, while implemented algorithms exist for this general problem, they were unable to give a solution to the system from contrast optimization. We proposed a strategy taking advantage of the determinantal structure of the system, in order to spread the computations over several smaller problems within the reach of existing algorithms. Using incidence varieties to model the rank defects, the regularity of the Gröbner basis algorithms used for computing eliminations is improved.

3.1. Weighted-homogeneous systems

Computational strategy Weighted homogeneous systems are a generalization of homogeneous systems. Given a *system of weights* $W = (w_1, \ldots, w_n) \in \mathbb{N}^n_{>0}$, the W-degree of the

monomial $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ is defined as

$$\deg_W \left(X_1^{\alpha_1} \cdots X_n^{\alpha_n}\right) = \sum_{i=1}^{n} w_i \alpha_i.$$

A polynomial is said to be W-homogeneous if all its monomials have the same W-degree.

Equivalently, a polynomial f is W-homogeneous with W-degree d if and only if

$$\hom_W(f) = f\left(X_1^{w_1}, \ldots, X_n^{w_n}\right)$$

is homogeneous with degree d.

This morphism gives a natural strategy for computing Gröbner bases for a weighted homogeneous system F: we can run the usual algorithms on $\hom_W(F)$. Some algorithms, such as F_4 or F_5, appears to become faster with this strategy, but algorithm FGLM becomes a bottleneck.

This strategy is also interesting from a theoretical point of view. Homogeneity is certainly the most studied structure with respect to Gröbner basis computations: this includes the complexity studies given above, but also detailed characterizations of the Hilbert series of an homogeneous ideal [Mor96], thin-grained complexity bounds under stronger genericity hypotheses [BFS14], and complexity analyses for overdetermined systems under a property conjectured to be generic [BFS04].

The question raised by all these results for homogeneous systems is how far these bounds and results can be transposed to the weighted homogeneous case.

It turns out that for a generic W-homogeneous system F, $\hom_W(F)$ is "sufficiently generic" as a homogeneous system, and a run of Algorithm Matrix-F_5 on $\hom_W(F)$ computes a Gröbner basis without any reduction to zero. In this case, as in the homogeneous case, its complexity can be determined in terms of the size of the computed matrices and the degree of regularity of the system.

In applications, systems are rarely weighted homogeneous. In this situation, as in the homogeneous case, the behavior of the algorithms is tied to reductions to zero of the highest W-degree components of the systems, and so, to whether these components satisfy regularity properties. If this condition is not satisfied, the behavior of algorithm F_5 is expected to be irregular, with reductions to zero and degree falls, and generic complexity analyses no longer apply: the algorithms perform worse in practice, and theory fails to quantify exactly how much worse.

This also explains why transforming a weighted homogeneous system into a homogeneous one yields significant improvements in practice. Indeed, given a generic weighted homogeneous systems, its highest degree components are unlikely to be large, and thus to satisfy regularity properties: informally, the smaller a class of polynomial systems is, the less likely there will be enough room for interesting generic properties to be satisfied.

We now turn to algorithm FGLM, for zero-dimensional systems: its complexity is bounded in terms of the number of solutions of the system, regardless of any generic property. The morphism $\hom_W$ multiplies this number of solutions by $\prod w_i$: for example consider $F = (X^2 - Y, X - 1)$, with the weights $(1, 2)$; its only solution is $(X = 1, Y = 1)$, but $\hom_W(F) = (X^2 - Y^2, X - 1)$ has 2 solutions: $(X = 1, Y = 1)$ and $(X = 1, Y = -1)$.

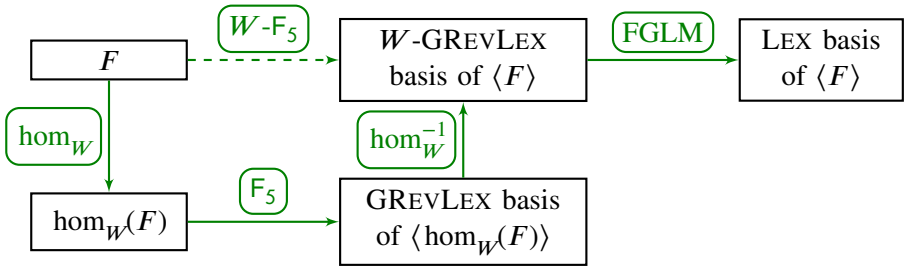

Figure 4: Algorithmic strategy for a zero-dimensional system, using the W-degree

This explains why algorithm FGLM becomes a bottleneck when computing a lexicographical Gröbner basis through $\hom_W$: the transformation artificially demultiplies the number of solutions, leading the algorithms to build and reduce unnecessarily large matrices.

On the other hand, it turns out that applying $\hom_W^{-1}$ to a reduced Gröbner basis of $\hom_W(F)$ yields a Gröbner basis for a weighted equivalent of the GRevLex order. This basis can also be computed directly from the original system by using the weighted degree instead of the total degree in algorithm F_4 or F_5.

This new basis can be used as input for algorithm FGLM, and it generates the same ideal as the original system F, hence it has the correct degree. Algorithm FGLM can compute a lexicographical basis of the ideal from this basis, with the same complexity bounds as above in terms of the degree of the ideal. The complete computational strategy is summed up on Figure 4.

Regularity properties and complexity bounds The definition of regular sequences does not involve the gradation, so it extends to a weighted setting. Regarding inhomogeneous systems, as in the total degree case, we consider whether the highest W-degree components form a regular sequence.

These properties are still generic amongst sequences of polynomials of given W-degree, as long as there exists at least one such sequence. It is proved straightforwardly by transposing the genericity proofs in the homogeneous case. However, the existence hypothesis is a necessary one, conditioning the system of weights: take for example $W = (2, 3)$ and $D = (5, 5)$, the only monomial of W-degree 5 being XY, so for any weighted-homogeneous system with respective W-degree D, the set of solutions is the union of $\{X = 0\}$ and $\{Y = 0\}$, and it is not a finite set.

We shall give hypotheses on the system of degrees ensuring that this hypothesis is satisfied. For example, if all degrees are divisible by the least common multiple of the weights, or if each degree d_i is divisible by the corresponding weight w_i, then regular sequences exist, and thus are generic. But these hypotheses are not necessary: take for example $W = (2, 3)$ and $D = (5, 6)$, the sequence $(XY, X^3 + Y^2)$ is regular.

Some results about weighted homogeneous systems were known before this thesis. Most notably, Bézout's bound giving the number of solutions of a zero-dimensional ideal admits a weighted version [Spa12, Th. 1.67 and 1.68]:

$$\deg(I) = \frac{\prod_{i=1}^{n} d_i}{\prod_{i=1}^{n} w_i},$$

which gives us immediately the complexity of algorithm FGLM in the strategy described above, for a square system defined by a regular sequence:

$$C_{\mathsf{FGLM}} = O\left(n\left(\frac{\prod_{i=1}^{n} d_i}{\prod_{i=1}^{n} w_i}\right)^3\right).$$

For a run of Algorithm Matrix-F_5 on $\hom_W(F)$, the usual complexity bounds from the homogeneous case can be applied, but they can actually be improved. The main asymptotical improvement comes from the work of combinatoricians on the *Sylvester denumerant* of d, that is the number of monomials with a given W-degree d. In particular, it is known [FS09] that asymptotically, this number of monomials is the number of monomials with *total* degree d, divided by the product of the weights. As such, the complexity bound for Algorithm Matrix-F_5 run on $\hom_W(F)$ is divided by $(\prod w_i)^3$ when compared to the original bound for homogeneous systems, given on page 13.

The bound on the degree of regularity can also be improved, with sufficient hypotheses on the system. Assuming that the system form a regular sequence, a first weighted version of Macaulay's bound yields that

$$d_{\text{reg}} \leq \sum_{i=1}^{n}(d_i - w_i) + \max\{w_i\}.$$

In other words, Algorithm Matrix-F_5 may skip the last

$$\sum_{i=1}^{n}(w_i - 1) - (\max\{w_i\} - 1)$$

degree steps in a run on $\hom_W(F)$.

This bound can be refined, but we need some stronger genericity hypothesis, containing some information about which variables are involved at each step of the computation. More precisely, we say that F is in *simultaneous Noether position* if the sequences $(f_1, X_2, \ldots, X_n)$, $(f_1, f_2, X_3, \ldots, X_n)$, …, $(f_1, \ldots, f_n)$ are all regular. Under this hypothesis, the following bound applies:

$$d_{\text{reg}} \leq \sum_{i=1}^{n}(d_i - w_i) + w_n.$$

In particular, when applicable, it implies that for the best complexity, one should order the variables such that the smallest variable (for the monomial order) is the one with the smallest weight.

This bound is still not sharp in full generality. We conjecture that with no hypothesis on the weights and W-degrees as long as there exist regular sequences, the following bound is true and optimal:

$$d_{\text{reg}} = w_n\left\lceil\frac{\sum_{i=1}^{n}(d_i - w_i) - g}{w_n}\right\rceil \tag{1}$$

with g is the *Frobenius number* of $(w_1, \ldots, w_n)$, that is the highest W-degree at which there exist no monomial (with the convention that $g = -1$ if there always exists a monomial, that is if $w_i = 1$ for some i).

Figure 5: Contrast optimization: both pictures were taken using the same setup of two tubes, the inner one containing deoxygenated blood and the outer one containing oxygenated blood. The left-hand picture is the reference picture, with no contrast optimization; the right-hand picture is contrast-optimized by saturating the inner sample.

Examples This structure appears in many applications, we already gave the example of the work of the authors of [Fau+13] on the DLP for Edwards elliptic curves.

Another example where weighted homogeneity appears naturally is polynomial inversion systems. Recall that given a set of polynomials $f_1, \ldots, f_n$ in $X_1, \ldots, X_n$, we may obtain their relations by adding variables $F_1, \ldots, F_n$ to the algebra, and eliminating $X_1, \ldots, X_n$ from $\langle F_1 - f_1, \ldots, F_n - f_n \rangle$. There is a natural system of weights to try when working with such systems, as first described in [Tra96]: if we give the F_i's a high-enough weight, we can ensure that the highest W-degree components of the system form a regular sequence and a sequence in Noether position. Furthermore, if the f_i have additional structure, for example if they are homogeneous, using the degree of f_i as the weight of F_i makes the highest W-degree components even larger, thus improving the regularity of the behavior of Algorithm F_5.

In Section 3.6, we shall present benchmarks computing the relations between fundamental invariants of the cyclic group or the dihedral group, between monomials, and between minors of matrices. We could obtain diverse speed-ups for the first step of the computational strategy (computing a GRevLex basis) for these systems, from a modest 1.5 to 4 for matrix minors to 10^5 for monomials.

3.2. Real root classification for determinantal systems Application to contrast optimization

Context The second class of structured systems that we studied arised in an application in medical imagery. Nuclear Magnetic Resonance (MNR) is a technique relying on applying a magnetic field to a body, and measuring the response of different matters. Contrast optimization is the problem of ensuring that two biological matters that we wish to distinguish on a picture, for example blood and water, stand apart: ideally, one should be completely black, and the other as bright as possible.

The saturation method is a contrast optimization technique, consisting of making the magnetic field variable in order to progressively obtain the optimal contrast. In order to find the optimal "path" towards the saturated state, control theory has been involved [Lap+12].

This strategy involves alternating so-called *bang-arcs* and *singular-arcs*, and estimating the

complexity of the path requires to estimate the number of switching points, in terms of the matters that we want to study. These switching points are characterized as the singularities of the determinant of a matrix whose entries are polynomials in the parameters and some variables, together with some polynomial inequalities [Bon+13].

This problem has two forms: in its most general form, both matters are free, each matter being described by two parameters. The structure of the system allows to eliminate one parameter, and we need to identify parts of $\mathbb{R}^3$ where the number of singularities does not change. A useful easier case is the case where one of the matters is water, leaving only 2 parameters.

This is a particular case of a more general problem: given a matrix $M(\mathbf{X}, \mathbf{G})$ whose entries are polynomial in some variables $\mathbf{X}$ and some parameters $\mathbf{G}$, and a target rank r, let V_r be the locus at which M has rank at most r. We wish to identify a dense subset U of the parameter space, together with a covering of U with open subsets, such that on each open subset, the number of singularities of V_r in the fiber over the parameters is constant.

In the case of the application, the physics of the system brings an additional constraint on the solutions: they have to be in the *Bloch ball*, which is described by inequations. Thus the problem is a real roots classification problem on a semi-algebraic set.

Algorithmic strategy The problem is a real roots classification problem for a semi-algebraic set described by a parameterized determinantal system and some inequalities.

The Cylindrical Algebraic Decomposition (CAD) method can be used to solve this problem, but none of its implementations can attack the application directly. On the other hand, it is known that adding equational constraints can dramatically improve the complexity of CAD computations. In our context, these equations should be the equations of the borders of the subdivision areas in the parameter space. This is the crux of real roots classification algorithms such as those presented in [LR07; YHX01]. However, none of the existing implementations of these algorithms for this problem were able to tackle the general case of the imagery problem.

We propose a strategy taking advantage of the determinantal structure of the varieties at hand, to guide existing algorithms towards a solution.

Under some hypotheses satisfied by the application, these strategies classify the cardinality of the fibers by identifying critical points of the projections, that is points where several single roots merge into one multiple root, and points where the variety meets the boundary of the relevant domain. In other words, we need to compute the singular locus of a determinantal variety, and then the singularities of this singular locus, the critical points of a projection restricted to it, and its intersection with the boundary of the domain.

The classical tool for computing critical and singular points is the Jacobian criterion, which characterizes singularities in terms of minors of the Jacobian matrix of the system defining the variety. In other words, a singular locus is again a determinantal variety.

Furthermore, the singular locus of a determinantal ideal is closely related to smaller minors of the same matrix. In particular, points where the rank of the matrix is $r-1$ are always singular points of the locus where the rank is r.

This is the first property that we use to steer the computations: we compute a *rank stratification* of the solutions, by separating those singularities where the matrix has exactly rank r (which we can compute using the Jacobian criterion), and those points where the matrix has rank at

most $r-1$, which are automatically singularities of the ideal.

Furthermore, recall that the classification strategy aims at computing a set of hypersurfaces of the parameter space $\mathbb{R}^t$, splitting it into open sets over which the number of singularities is constant. So if we know that some component C of the singular locus of the determinantal variety has dimension at most $t-1$, we only need to compute the equation of one hypersurface containing its image by the projection on the parameter space: this hypersurface will necessarily contain the projection of the singular points of C, the critical values of the projection restricted to C and the intersection of C with the boundaries of the domain.

Under some generic hypotheses, the dimension of rank strata is determined by the target rank defect and the number of variables and parameters. It turns out that the high rank stratum necessarily has dimension less than t, thus reducing the problem to a single elimination for that stratum.

The last regularization step is the use of incidence variety to model singularities of the determinants where needed. Algorithmically, we add $k(k-r)$ variables to the system, build a matrix N with k rows and $k-r$ columns, each entry being one of the new variables, and we build the system with the entries of $M \cdot N$. In order to ensure that the resulting kernel vectors are linearly independent, we pick at random a matrix U with $k-r$ rows and k columns, and we add to the system the entries of $U \cdot N - \mathrm{Id}_{k-r}$.

Results for the application This strategy was used for answering questions related to the contrast imaging problem.

In particular, an open question for the case of water was whether the number of singularities was a global invariant, not depending on the second matter: for all experimental samples, there was only one such singularity.

It turns out that the answer is *no*: we identified 3 areas of the parameter space where the number of singularities is 1, 2 and 3 respectively. See Figure 6: the red area is out of the physically relevant domain; for parameters in the areas containing blue diamonds, there are two singularities in the fiber; for parameters in the areas containing green circles, there are three singularities in the fiber; and for all other areas, there is only one singularity.

These results can be obtained in about 10 s using the new strategy, while a direct approach using Gröbner bases requires 100 s. Furthermore, we were able to use this strategy to obtain the boundaries of the classification in the general case. This computation takes less than 2 h, while it is intractable for the general strategies.

4. Perspectives

In this section, we list some questions left open by this thesis, as well as potential extensions and generalizations that we would find interesting.

Open questions and conjectures for W-homogeneous systems First, there is the problem of W-compatibility: what are the systems of degrees and weights such that there exists W-homogeneous regular sequences, and thus that they are generic? We were able to give sufficient conditions, but counter-examples show that they are not necessary.

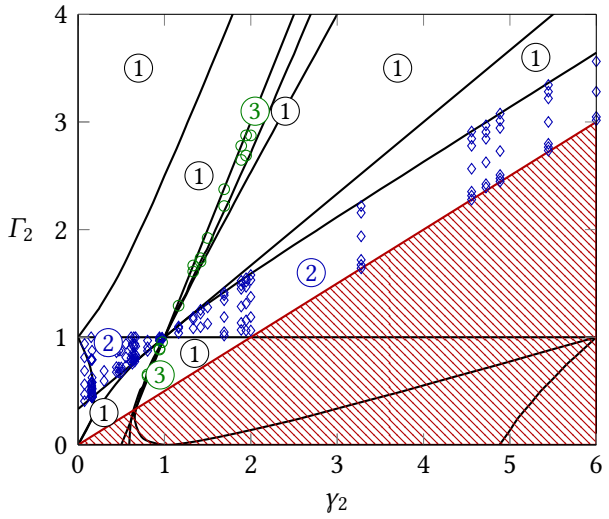

Figure 6: Decomposition of the parameter space in the case of water: the parameters are the relaxation times for the second matter, γ_2 and Γ_2, and in each area delimited by the black curves, the cardinality of the fiber of the singularities of V_r is constant and written as a circled number; the red area corresponds to physically irrelevant parameters.

Genericity results are useful because testing whether a given system is a regular sequence is not easy in general: if one knows that most sequences similar to the system at hand are regular, it makes sense to run the algorithms assuming that the hypothesis is satisfied.

As of now, for systems of degrees and weights falling out of the sufficient conditions for regular sequences to exist, this assumption is riskier. Obtaining a set of necessary and sufficient conditions could complete the characterization of the genericity of weighted homogeneous regular sequences and make this procedure more likely to succeed.

Second, we conjectured formula (1) as a sharp bound of the degree of regularity of a sequence in simultaneous Noether position, without any hypothesis on the weights. So far, we have not been able to prove it.

And last, it would be interesting to be able to characterize semi-regular W-homogeneous sequences in terms of their Hilbert series, once again without any hypothesis on the weights.

Several systems of weights Weighted homogeneous systems have been studied considering only one system of weights. However, some systems may be W-homogeneous for several systems of weights W. Consider for example the polynomial

$$X^3Y^2Z + X^4Z^2 + X^2Y^4,$$

it is $(1, 1, 1)$-homogeneous with degree 6, and $(1, 2, 3)$-homogeneous with degree 10.

This new constraint gives extra structure to the Macaulay matrix of the system. Exploiting this structure should yield further improvements on the complexity of Matrix-F_5 for these systems.

Furthermore, polynomials with are W_i-homogeneous for all systems of weights W_i in some family are also W-homogeneous for any W obtained by linear combination of the W_i's: for example, the above polynomial is also $(2, 3, 4)$-homogeneous with degree 16. It may be interesting to see if this property can be used to find systems of weights leading to easier computations.

Non-positive weights An extension of the previous point could be to consider systems of weights including weights set to zero or a negative integer. With only one system of weights, this leads to problems because homogeneous components do not necessarily have finite dimension. However, with several systems of weights, it is possible to ensure that this does not happen, and it would be useful to know what algorithmic strategies we can use in this situation.

This class of systems is very wide, it would for example include multi-homogeneous systems (systems where the variables are split into groups, and polynomials are homogeneous in total degree for each group): for example, consider a polynomial f in $\mathbb{K}[X_1, X_2, Y_1, Y_2]$, and assume that it is bilinear for the variables X_1, X_2 and Y_1, Y_2. In terms of weights, consider the systems of weights $W_1 = (1, 1, 0, 0)$ and $W_2 = (0, 0, 1, 1)$; f being bilinear means that f is W_1-homogeneous with W_1-degree 1 and W_2-homogeneous with W_2-degree 1.

Weight search Returning to the case of only one system of weights, until now, we have only worked where systems for which the user was able to guess a good system of weights (either because it was natural in the application, or through trial and error). It would be interesting to describe a strategy, or at least a good set of heuristics, giving good candidate systems of weights for an arbitrary inhomogeneous system.

Testing that a system is a regular sequence is not computationally easy, so using that as a criterion is impractical. On the other hand, we could look for systems of weights making the highest W-degree component large-enough for it to be generically regular.

Experimentally, we also observed systems which behaved better with a system of weights such that the highest W-degree component is small, but with a very large W-homogeneous component at smaller W-degree. Characterizing these situations could lead to further improvements on a strategy for picking up good systems of weights.

Application to imagery Our work on real roots classification for the imagery problem also left several open questions and possible extensions.

First, we only described algorithms, and showed experimental data confirming that they are more efficient. It would be interesting to complete the analysis with complexity bounds for the whole process.

Furthermore, the algorithmic strategy described for the roots classification problem for the contrast imagery problem builds a stratification by rank. To the best of our knowledge, the question of whether this stratification is physically meaningful has never been raised.

As for the extensions, the root classification problem is only one of the many computational problems posed by specialists in control theory working on constrast optimization for the MNR. For example, they would be interested in a rational parametrisation of the roots, or in a classification of the parameter space according to the number of connected components in the fibers for higher-dimensional varieties. All these questions offer new computational challenges.

Part I

Preliminaries

Chapter 1

Algebra and geometry

In this chapter, we recall concepts of commutative algebra and algebraic geometry that will be used in the rest of the thesis. In Section 1.1, we give some basic definitions about rings and ideals, and in particular the properties of gradations on polynomial algebras. In Section 1.2, we focus on some properties of polynomial systems, which will be useful in the context of Gröbner bases. In Section 1.3, we briefly recall the correspondence between algebraic and geometric notions, and in particular we give an interpretation of the previous properties. In Section 1.4, we examine the genericity of regularity properties. Finally, in Section 1.5, we give some definitions and properties of varieties characterizing low-rank matrices.

This chapter and the next one do not contain any original contribution. For most results, we give a reference to a point in the literature where a proof may be found. We only give an explicit proof for folklore results, or when the proof itself is useful for subsequent chapters (*e.g.* if it is generalized later).

1.1. Commutative algebra

1.1.1. Ideals

We assume that classic definitions about rings, algebras and ideals are known. We shall consider rings and algebras, which we will suppose to be commutative and with a unity.

Definition 1.1. Let R be a ring. A *divisor of zero* is an element $x \in R$, non-zero, such that

$$\exists y \in R, y \neq 0, xy = 0.$$

Definition 1.2 (Noetherian ring)**.** A ring R is *noetherian* if and only if any increasing chain of ideals

$$I_0 \subseteq I_1 \subseteq \cdots \subseteq I_n \subseteq \ldots$$

is stationary.

This property is sometimes defined using the following classic characterization:

Proposition 1.3 ([Eis95, Sec. 1.4])**.** *A ring R is noetherian if and only if all ideals of R are finitely generated.*

Proof. Let R be a noetherian ring and I an ideal of R. If I is not finitely generated, then we can find a family of polynomials $(f_i)_{i\in\mathbb{N}} \subset I$ such that no finite subset of this family generates I. Let $I_i = \langle f_0, \ldots, f_i \rangle$, the sequence of ideals $(I_i)_{i\in\mathbb{N}}$ is strictly increasing, which is impossible since R is noetherian.

Conversely, let R be a ring such that all ideals of R are finitely generated. Let $I_0 \subseteq I_1 \subseteq \cdots \subseteq I_n \subseteq \ldots$ be an increasing chain of ideals, the limit $\bigcup_{i\in\mathbb{N}} I_i$ is an ideal, and by hypothesis it is finitely generated. All of its generators necessarily belong to one of the ideals in the chain, and the sequence stations at that point. □

1.1.2. Graded rings

Definition 1.4 (Gradation on a ring, degree, homogeneous components)**.** Let R be a ring (*resp.* algebra), R is *graded* if and only if there exists a decomposition (called a *gradation* or *grading* of R) into additive groups (*resp.* vector spaces)

$$R = \bigoplus_{i=0}^{\infty} R_i$$

such that $R_i \cdot R_j \subset R_{i+j}$: the product of elements of degree i and j has degree $i + j$. We say that the elements of R_i are *homogeneous of degree i*.

Let $p \in R$, there exists $d \in \mathbb{N}$ and $p_i \in R_i$ ($i \in \{0, \ldots, d\}$) such that

$$p = \sum_{i=0}^{d} p_i \text{ and } p_d \neq 0.$$

We then say that p has *degree d*, and p_i is called the *homogeneous component of p of degree i*. If all p_i's are zero except for one, p is *homogeneous*.

Definition 1.5 (Homogeneous ideal)**.** Let R be a graded ring, and I an ideal of R. We say that I is *homogeneous* if it satisfies one of the following equivalent conditions:

1. I is generated by homogeneous elements
2. For any $p \in I$, all homogeneous components of p belong to I

Definition 1.6 (Graded morphism)**.** Let R and S be two graded rings. A morphism $\varphi : R \longrightarrow S$ is said to be *graded* if it sends homogeneous elements of R onto homogeneous elements of S:

$$\forall d \in \mathbb{N}, \exists e \in \mathbb{N}, \varphi(R_d) \subset S_e$$

1.1.3. Polynomial algebras

Definition 1.7. Let $\mathbb{K}$ be a field and n a positive integer, let A be the *polynomial algebra* $\mathbb{K}[\mathbf{X}]$ with coefficients in $\mathbb{K}$ and n *indeterminates* $\mathbf{X} = X_1, \ldots, X_n$.

A product of indeterminates $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ is a *monomial*, and the integer vector $(\alpha_1, \ldots, \alpha_n)$ is its *exponent*.

A *term* is the product of a coefficient and a monomial.

In particular, monomials generate a polynomial algebra as a $\mathbb{K}$-vector space, and terms generate a polynomial algebra as an additive group.

We shall use the same notations throughout this section.

Theorem 1.8. *The polynomial algebra A is a noetherian ring.*

This is a consequence of a more general result:

Theorem 1.9 (Hilbert's basis theorem, [Eis95, Th. 1.2])**.** *Let R be a noetherian ring, then the polynomial algebra $R[X]$ is noetherian.*

Remark 1.10. Theorem 1.8 for polynomial rings admits a more elementary proof than the more general Theorem 1.9, see for example [CLO07, Ch. 2, sec. 5, th. 4] or [Eis95, Ex. 15.15].

An important feature of polynomial algebras is that they can be graded.

Definition 1.11 (Total degree, weighted degree)**.** Let $W = (w_1, \dots, w_n) \in \mathbb{N}^n$ be a family of integers. We call W a *system of weights*, and define the *W-degree* (or *weighted degree*) of a monomial as

$$\deg_W\left(\prod_{i=1}^{n} X_i^{\alpha_i}\right) = \sum_{i=1}^{n} w_i \alpha_i,$$

and let $A_{W,d}$ be the vector space generated by all monomials m such that $\deg_W(m) = d$. Polynomials in $A_{W,d}$ are called *W-homogeneous* (or *weighted homogeneous*).

In the special case $W = (1, \dots, 1)$, the W-degree is called *total degree*, or simply *degree*. In this case, W-homogeneous polynomials are simply called *homogeneous*.

Proposition 1.12. *Let $A = \bigoplus A_d$ be a gradation such that monomials are homogeneous. Then there exists $W \in \mathbb{N}^n$ such that*

$$\forall d \in \mathbb{N}, A_d = A_{W,d}$$

Proof. Consider a degree function deg on A. For all $i \in \{1, \dots, n\}$, let $w_i = \deg(X_i)$. The multiplicative property of the degree implies that deg is $\deg_{(w_1,\dots,w_n)}$. □

Throughout this thesis, we will only consider non-degenerate systems of weights, that is systems of weights such that $w_i > 0$ for all i.

An important concept is that of homogeneization:

Definition 1.13 (Homogenization of polynomials)**.** Let $A_h = \mathbb{K}[H, \mathbf{X}]$. The *homogeneization morphism* is defined as

$$\begin{array}{rccl} \bullet^h : & A & \longrightarrow & A_h \\ & f & \longmapsto & H^d f\left(\frac{X_1}{H}, \dots, \frac{X_n}{H}\right) \text{ where } d \text{ is the degree of } f \end{array}$$

The *dehomogenization morphism* is defined as

$$\begin{array}{rccl} \bullet^a : & A_h & \longrightarrow & A \\ & f & \longmapsto & f(H = 1) \end{array}$$

It has the following immediate properties:

Proposition 1.14.

- *The homogenization morphism is graded and sends polynomials of degree d onto polynomials of degree d;*
- *For any polynomial f in A of degree d (not necessarily homogeneous), f^h is homogeneous of degree d;*
- *For any $f \in A$, $(f^h)^a = f$;*
- *For any $f_h \in A_h$, if f_h is not divisible by H, $(f_h^a)^h = f_h$.*

Definition 1.15 (Homogenization of ideals)**.** Let I be an ideal of A, the homogenization ideal I^h is defined as

$$I^h = \langle f^h \mid f \in I \rangle;$$

let J be an ideal of A_h, the dehomogenization ideal I^a is defined as

$$I^a = \langle f^a \mid f \in J \rangle.$$

The element-wise properties above have the following consequences on ideals:

Proposition 1.16. *Let I be an ideal of A and J be an homogeneous ideal of A_h. Then:*

- *I^h is an homogeneous ideal of A_h;*
- *$(I^h)^a = I$;*
- *$(J^a)^h = J$.*

1.1.4. Combinatorics of monomials

Let $d \in \mathbb{N}$, the homogeneous component A_d forms a finite dimensional $\mathbb{K}$-vector space, and its dimension is the number of monomials at degree d.

Counting these monomials in the total degree case is a classic combinatorial exercise:

Proposition 1.17. *Let $n \in \mathbb{N}$ and $d \in \mathbb{N}$. The number of monomials in n variables having total degree d is*

$$N_{n,d} = \binom{n+d-1}{d}.$$

The number of monomials in n variables having total degree at most d is

$$N'_{n,d} = N_{n+1,d} = \binom{n+d}{d}.$$

In the more general case, this number is called a denumerant, as introduced by Sylvester:

Definition 1.18 (Sylvester denumerant, [SF82])**.** Let W be a set of integers and $d \in \mathbb{N}$, the *denumerant* $N_{W,d}$ is the number of non-negative integer solutions $(\alpha_1, \dots, \alpha_n)$ to the equation

$$w_1\alpha_1 + \cdots + w_n\alpha_n = d.$$

There are formulas expressing the denumerants for some specific weights W [Com74, Sec. 2.6], but in the most general case, there is no known formula expressing a denumerant as a function of W and d. The asymptotic behavior of these denumerants is well-known:

Proposition 1.19 ([FS09, Prop. IV.2])**.** *If W is fixed and d tends towards infinity, then*

$$N_{W,d} \sim \frac{\gcd(W)}{\Pi(W)} N_{n,d} = \frac{\gcd(W)}{\Pi(W)} \binom{n+d-1}{d}.$$

Additionally, some bounds are known:

Proposition 1.20 ([Agno2, Th. 3.3 and 3.4])**.** *Define*

$$\begin{aligned} A &:= \sum_{i=2}^{n} w_i \frac{\gcd(w_1, \dots, w_i)}{\gcd(w_1, \dots, w_{i-1})} \\ B &:= \sum_{i=2}^{n} w_i \left(\frac{\gcd(w_1, \dots, w_i)}{\gcd(w_1, \dots, w_{i-1})} - 1 \right) - n \\ \gcd(W) &:= \gcd(w_1, \dots, w_n) \\ \Pi(W) &:= w_1 \cdots w_n, \end{aligned}$$

where $\gcd(a_1, \dots, a_k)$ denotes the greatest common divisor of the integers $a_1, \dots, a_k$. Then the number $N_{W,d}$ of monomials with W-degree d is bounded by

$$\frac{\gcd(W)}{\Pi(W)} N_{n,d-B+1} \leq N_{W,d} \leq \frac{\gcd(W)}{\Pi(W)} N_{n,d+A-n+1}.$$

Another interesting combinatorial number related to sums of integers is the Frobenius number [Alf05, Ch. 1].

Definition 1.21 (Frobenius number)**.** Let W be a set of integers, assume that they are coprime. The *Frobenius number* of W is the largest integer d such that the equation

$$w_1\alpha_1 + \cdots + w_n\alpha_n = d.$$

has no non-negative integer solutions $(\alpha_1, \dots, \alpha_n)$.

Remark 1.22. Algebraically, it corresponds to the largest W-degree d such that there exists no monomials with W-degree d, or equivalently, such that the homogeneous component of W-degree d in A is 0.

Remark 1.23. Computing this number is in general a hard computational problem: it has been proved to be NP-hard [Alf05, Th. 1.3.1]. In particular, there is no general closed-form formula if $n > 2$ ([Alf05, Sec. 1.3]).

Remark 1.24. An easy case is if one of the integers is 1, because then the equation always have solutions, and the Frobenius number does not exist.

Remark 1.25. The Frobenius number is sometimes called *coin number*, because of its application to the problem of making change in small coins. Another example of application comes from restauration: the *McNuggets problem* is deciding, given an integer d, whether it is possible to obtain d chicken nuggets using boxes of 6, 9 and 20. If it is the case, d is called a *McNuggets number*, and the Frobenius number of $\{6, 9, 20\}$ is the largest non-McNuggets number ([Alf05, Ch. 1]).

1.1.5. Hilbert series

In the previous section, we saw that computing the number of monomials at a given W-degree directly is a non-trivial problem. However, results such as the asymptotic equivalent of the denumerants (Prop. 1.19) have been known for a long time. These results were proved by considering generating series, instead of individual denumerants.

Definition 1.26 (Generating series)**.** Let $f : \mathbb{N} \longrightarrow \mathbb{N}$ be an integer function. The *generating series* of f is the formal series

$$\sum_{d=0}^{\infty} f(d)T^d \in \mathbb{N}[[T]].$$

Generating series are useful because many usual combinatorial operations can be converted into algebraic operations on the generating series. For the case at point, the generating series for denumerants is well known:

Proposition 1.27 ([Com74, Sec. 2.6, Th. A], [FS09, Prop. I.1])**.** *Let W be a set of integers, the generating series for the denumerants $N_{W,\bullet}$ is given by*

$$\sum_{d=0}^{\infty} N_{W,d}T^d = \frac{1}{\prod_{i=1}^{n}(1 - T^{w_i})}.$$

This generating series is a particular case of a Hilbert series:

Definition 1.28 ([Eis95, Ex. 12.12])**.** Let R be a graded $\mathbb{K}$-algebra such that all homogeneous components, as $\mathbb{K}$-vector spaces, have finite dimension. The *Hilbert series* of R is the power series defined by

$$\mathrm{HS}_R(T) = \sum_{d=0}^{\infty} \dim_{\mathbb{K}}(R_d)T^d \in \mathbb{N}[[T]].$$

Proposition 1.29. *The Hilbert series of A, graded with respect to a system of weights W, is*

$$\mathrm{HS}_A(T) = \frac{1}{\prod_{i=1}^{n}(1 - T^{w_i})}.$$

Hilbert series of quotients A/I where I is a homogeneous ideal will be very useful, because many interesting parameters of the ideal can be read from the series. In section 1.2, we shall see some examples of polynomial systems whose Hilbert series can be computed very easily.

Proposition 1.30 ([Eis95, Ex. 12.12])**.** *Let $I \subset A$ be a W-homogeneous ideal. Then there exist a polynomial $P \in \mathbb{Z}[T]$, and an integer $r \in \mathbb{N}$, such that the Hilbert series of A/I can be written*

$$\mathrm{HS}_{A/I}(T) = \frac{P(T)}{\prod_{i=1}^{n}(1 - T^{w_i})}.$$

Remark 1.31. This result will be proved as a by-product of the proof of Prop. 1.38 in Section 1.2.1.

Definition 1.32. With the same notations:

- the *codimension* of I is the multiplicity r of 1 as a root of P;
- the *dimension* of I is $d = \dim(I) = n - r$;
- the *index of regularity* of I is $i_{\mathrm{reg}}(I) = \deg(P) - r$;
- if I is homogeneous for the total degree, the *degree* of I is $\deg(I) = Q(1)$ where $Q = \mathrm{HS}_{A/I}(T) \cdot (1 - T)^d$.

Remark 1.33. In Section 1.3.4, we shall see that the dimension and degree have geometrical interpretations. For example, if $r = n$, the dimension of I is 0, and $\mathrm{HS}_{A/I}$ is a polynomial. The degree of I is then the dimension of A/I as a $\mathbb{K}$-vector space. In this case, there are only a finite number of zeroes common to all polynomials of I, and this number, counted with multiplicities, is $\deg(I)$.

Hilbert series (and dimension, codimension, etc.) can be defined for non-homogeneous ideals as well, by taking the Hilbert series of A^h/I^h as a definition.

To conclude this section, we give a last interpretation of the Hilbert series of a homogeneous quotient ring, which will be of immediate use when describing Gröbner basis algorithms.

Definition 1.34 (Macaulay matrix)**.** Let W be a system of weights, $F = (f_1, \ldots, f_m) \in \mathbb{K}[\mathbf{X}]^m$ a W-homogeneous polynomial system with respective W-degree $(d_1, \ldots, d_m)$, and $d \in \mathbb{N}$ a W-degree. For any $i \in \{1, \ldots, m\}$, let $\mathcal{M}_d$ be the set of monomials μ with W-degree d: The family $\mathcal{F}_d$ of all products μf_i's, $i \in \{1, \ldots, m\}$, $\mu \in \mathcal{M}_{d-d_i}$, is a linear basis of the homogeneous component of $\langle F \rangle$ at W-degree d. The *Macaulay matrix* of F at W-degree d is the matrix of $\mathcal{F}_d$ in the canonical monomial basis (whose elements are the monomials of W-degree d).

$$\mathrm{Mac}_d(F) = \begin{array}{c} \\ \mu_{1,1}f_1 \\ \vdots \\ \mu_{1,N_1}f_1 \\ \mu_{2,1}f_2 \\ \vdots \\ \vdots \\ \mu_{m,N_m}f_m \end{array} \begin{array}{c} \begin{array}{cccc} \mu_1 & \mu_2 & \cdots & \mu_N \end{array} \\ \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

Remark 1.35. Macaulay matrices are a generalization of Sylvester matrices ([GKZ94, Ch. 12]) to the multivariate case.

Proposition 1.36. *If F is a W-homogeneous polynomial system, the Hilbert series of $A/\langle F\rangle$ is the generating series of the corank of $\mathrm{Mac}_d(F)$, defined as the difference between its number of columns and its rank.*

1.2. Regularity properties

1.2.1. Regular sequences

Definition 1.37 (Regular sequences)**.** Let m be an integer, W a system of weights, and $F = (f_1, \dots, f_m) \in A^m$ a family of W-homogeneous polynomials. We say that F is a *regular sequence* if for any $i \in \{1, \dots, m\}$, f_i is not a zero-divisor in $A/\langle f_1, \dots, f_{i-1}\rangle$. In this case, we say that the quotient ring $A/\langle f\rangle$ is a *complete intersection.*

This notion has several equivalent characterizations:

Proposition 1.38 ([Bar04, Prop. 1.7.4(2)])**.** *Let m be an integer, W a system of weights, $F = (f_1, \dots, f_m) \in A^m$ a sequence of W-homogeneous polynomials, $(d_1, \dots, d_m)$ their respective W-degree, and $I = \langle F\rangle$. The Hilbert series of A/I satisfies the inequality (coefficient-wise)*

$$\mathrm{HS}_{A/I}(T) \geq \frac{\prod_{i=1}^{m}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})}$$

with equality if and only if F is a regular sequence.

Proof. For $i \in \{1, \dots, m\}$, let $I_i = \langle f_1, \dots, f_i\rangle$. We will prove the result by recurrence on i.

If $i = 0$, then $A/I = A$ and the result is a restatement of the formula for the Hilbert series of a polynomial algebra.

Let $i > 0$ and assume that the result is proved for I_{i-1}. The multiplication by f_i gives an exact sequence:

$$0 \longrightarrow K_i \longrightarrow A/I_{i-1} \xrightarrow{\cdot f_i} A/I_{i-1} \longrightarrow A/I_i \longrightarrow 0,$$

where K_i is the kernel of the application. For any $d \in \mathbb{N}$, this exact sequence restricted to W-homogeneous components of W-degree d is:

$$0 \longrightarrow (K_i)_d \longrightarrow (A/I_{i-1})_d \xrightarrow{\cdot f_i} (A/I_{i-1})_{d+d_i} \longrightarrow (A/I_i)_{d+d_i} \longrightarrow 0,$$

which implies for the dimensions:

$$\dim((A/I_i)_{d+d_i}) = \dim((A/I_{i-1})_{d+d_i}) - \dim((A/I_{i-1})_d) + \dim((K_i)_d).$$

Multiplying by T^d and taking the sum for d ranging over $\mathbb{N}$, we get

$$\mathrm{HS}_{A/I_i}(T) = (1 - T^{d_i})\mathrm{HS}_{A/I_{i-1}}(T) + \sum_{d=0}^{\infty} T^{d+d_i} \dim((K_i)_d).$$

By recurrence hypothesis,

$$\mathrm{HS}_{A/I_{i-1}}(T) = \frac{\prod_{j=1}^{i-1}(1-T^{d_j})}{\prod_{j=1}^{n}(1-T^{w_j})} + R_{i-1}(T),$$

where $R_{i-1}(T)$ is a series whose coefficients are all nonnegative. Hence

$$\mathrm{HS}_{A/I_i}(T) = \frac{\prod_{j=1}^{i}(1-T^{d_j})}{\prod_{j=1}^{n}(1-T^{w_j})} + R_{i-1}(T) + \sum_{d=0}^{\infty} T^{d+d_i} \dim((K_i)_d), \tag{1.1}$$

and all coefficients in the rightmost summand are nonnegative as well, so we get the wanted inequality.

Furthermore, if the sequence $(f_1, \dots, f_i)$ is regular, then $K_i = 0$ and by recurrence hypothesis, $R_{i-1}(T) = 0$, so there is indeed equality. Conversely, assume that

$$\mathrm{HS}_{A/I_i}(T) = \frac{\prod_{j=1}^{i}(1-T^{d_j})}{\prod_{j=1}^{n}(1-T^{w_j})}$$

From Equation (1.1), $R_{i-1}(T) = 0$ and $K_i = 0$. By recurrence hypothesis, if $R_{i-1}(T) = 0$, then $(f_1, \dots, f_{i-1})$ is a regular sequence; and by definition, if $K_i = 0$, then f_i is not a zero divisor in A/I_{i-1}, so $(f_1, \dots, f_i)$ is indeed regular. □

Corollary 1.39 ([Baro4, Prop. 1.7.4(3)]). *If F is a regular sequence, then any permutation of F is also regular.*

Proof. The characterization from Proposition 1.38 does not depend on the order of the polynomials. □

1.2.2. Description of the Hilbert series of a homogeneous regular sequence

In the total degree case, the Hilbert series of a zero-dimensional regular sequence admits the following description.

Theorem 1.40 ([Mor96, Prop. 2.2]). *Let $D = (d_1, \dots, d_n) \in \mathbb{N}^n$, and let a_d be the coefficient of degree d in the series*

$$S_D(T) = \frac{\prod_{i=1}^{n}(1-T^{d_i})}{(1-T)^n}.$$

Further let

- $\delta = \sum_{i=1}^{n}(d_i - 1)$;
- $\delta^* = \sum_{i=1}^{n-1}(d_i - 1) = \delta - d_n + 1$;
- $\sigma = \min(\delta^*, \lfloor \delta/2 \rfloor)$;
- $\mu = \delta - 2\sigma$.

Then:

- *the series $S_D(T)$ is a polynomial of degree δ*
- *it is self-reciprocal (i.e. for any d, $a_d = a_{\delta-d}$)*
- *the coefficients a_d satisfy:*

$$\begin{aligned} \forall d \in \{0, \dots, \sigma - 1\}, &\quad a_d < a_{d+1} \\ \forall d \in \{\sigma, \dots, \sigma + \mu - 1\}, &\quad a_d = a_{d+1} \\ \forall d \in \{\sigma + \mu, \dots, \delta\}, &\quad a_d > a_{d+1} \end{aligned}$$

Example 1.41. Take a homogeneous regular sequence of $A = \mathbb{K}[X, Y, Z]$ with respective total degree $(2, 4, 8)$: for example, $F = (X^2, Y^4, Z^8)$. The Hilbert series of $A/\langle F \rangle$ is

$$\begin{aligned} H(T) &= \frac{(1 - T^2)(1 - T^4)(1 - T^8)}{(1 - T)^3} \\ &= (1 + T)(1 + T + T^2 + T^3)(1 + T + \cdots + T^7) \\ &= 1 + 3T + 5T^2 + 7T^3 + 8T^4 + 8T^5 + 8T^6 + 8T^7 + 7T^8 + 5T^9 + 3T^{10} + T^{11}. \end{aligned}$$

The value of the coefficient of degree d in this series, as a function of d, is plotted in Figure 1.1.

In this case, the parameters from Theorem 1.40 are:

$$\begin{aligned} \delta &= (2 - 1) + (4 - 1) + (8 - 1) = 11 \\ \delta^* &= (2 - 1) + (4 - 1) = 4 \\ \sigma &= \min(4, 10) = 4 \\ \mu &= 11 - 2 \cdot 3 = 3. \end{aligned}$$

Indeed, the polynomial $H(T)$ is self-reciprocal (the plot is symmetric around $d = 5.5$) and its coefficients are increasing until degree σ, then stationary until degree $\sigma + \mu$, then decreasing.

In Section 3.2.5 we shall generalize this theorem in a weighted setting.

1.2.3. Noether position

We recall the following definition of ring theory:

Definition 1.42 (Integral element, integral morphism, integral extension)**.** Let $R \subset S$ be two rings, and $x \in S$. An element $x \in S$ is *integral* over R if there exists a *monic* polynomial with coefficients in R annihilating x:

$$\exists\, r_0, \dots, r_{d-1} \in R, x^d + r_{d-1}x^{d-1} + \cdots + r_0 = 0.$$

An *integral morphism* $\varphi : R \longrightarrow S$ is a morphism such that any element of S is integral over $\varphi(R)$. If the inclusion $R \hookrightarrow S$ is an integral morphism, then $R \subset S$ is an *integral extension.*

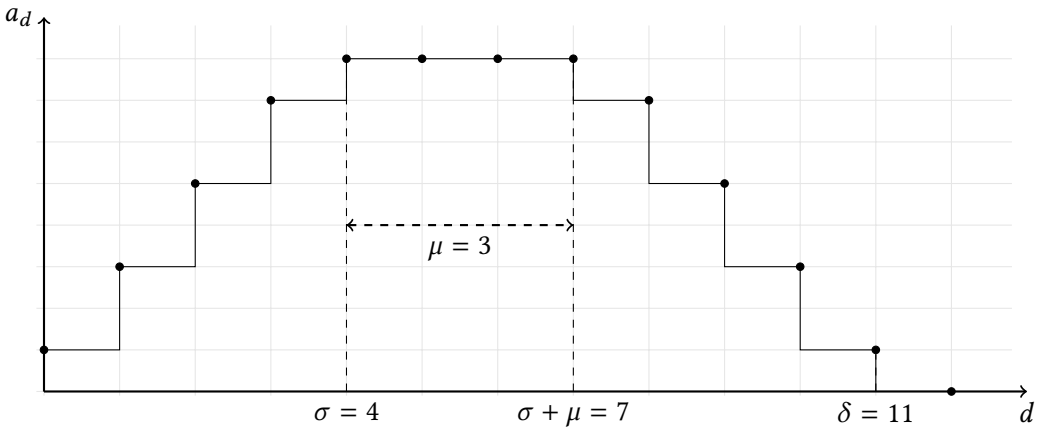

Figure 1.1: Shape of the Hilbert series of a homogeneous regular sequence with total degree $(2, 4, 8)$

Definition 1.43 (Noether position)**.** Let $F = (f_1, \ldots, f_m)$ be a sequence of polynomials in $\mathbb{K}[\mathbf{X}]$. The ideal $\langle F \rangle$ is said to be *in Noether position w.r.t.* the variables $X_1, \ldots, X_m$ if it satisfies the two following properties:

- for $i \in \{1, \ldots, m\}$, the canonical image of X_i in $\mathbb{K}[\mathbf{X}]/I$ is an algebraic integer over $\mathbb{K}[X_{m+1}, \ldots, X_n]$;
- $\mathbb{K}[X_{m+1}, \ldots, X_n] \cap I = 0$.

In other words, the morphism

$$\mathbb{K}[X_{m+1}, \ldots, X_n] \longrightarrow \mathbb{K}[X_1, \ldots, X_n]/\langle F \rangle$$

is integral and injective.

The system F is said to be *in simultaneous Noether position* (or in SNP) *w.r.t.* the order $X_1 > X_2 > \cdots > X_n$ if for any $1 \leq i \leq m$, the ideal $\langle f_1, \ldots, f_i \rangle$ is in Noether position *w.r.t.* the variables $X_1, \ldots, X_i$.

To conclude this section, we give several useful characterizations of Noether position. They appear to be folklore (see for example [BFS14, Prop. 6]), we give a proof for completeness.

For $m \in \{1, \ldots, n\}$, let θ_m be the morphism evaluating $X_{m+1}, \ldots, X_n$ to 0:

$$\begin{array}{rccccc} \theta_m & : & \mathbb{K}[\mathbf{X}] & \longrightarrow & \mathbb{K}[\mathbf{X}] & \\ & & X_i & \longmapsto & X_i & (i \leq m) \\ & & X_i & \longmapsto & 0 & (i > m) \end{array}$$

Proposition 1.44. *Let $m \leq n$, and let $F = (f_1, \ldots, f_m)$ be a sequence of polynomials. The following statements are equivalent:*

1. *the sequence F is in Noether position w.r.t. the variables $X_1, \ldots, X_m$;*

2. *the sequence* $F_{ext} \coloneqq (f_1, \ldots, f_m, X_{m+1}, \ldots, X_n)$ *is regular;*
3. *the sequence* $\theta_m(F) = F(X_1, \ldots, X_m, 0, \ldots, 0)$ *is in Noether position w.r.t. the variables* $X_1, \ldots, X_m$*;*
4. *the sequence* $\theta_m(F)$ *is regular.*

Proof. (1 $\implies$ 2). Let I be the ideal generated by F. In Section 1.3.5, we shall state a geometric interpretation of Noether position (Proposition 1.65), which shows that the canonical projection onto the $n - m$ last coordinates

$$\pi : V(I) \longrightarrow V(\langle X_1, \ldots, X_m \rangle)$$

is a surjective morphism with finite fibers. This implies that the variety $V(\langle F_{\text{ext}} \rangle) = \pi^{-1}(0)$ is zero-dimensional, and so the sequence is regular.

(2 $\implies$ 1). This statement will be proved using basic notions of Gröbner basis theory, which are introduced in Sections 2.1.1 and 2.1.2. Let $i \leq m$, we want to show that X_i is integral over the ring $\mathbb{K}[X_{m+1}, \ldots, X_n]$. Since F_{ext} defines a zero-dimensional ideal, there exists $n_i \in \mathbb{N}$ such that $X_i^{n_i} = \mathsf{LT}(f)$ with $f \in \langle F_{\text{ext}} \rangle$ for the Lex ordering with $X_n > \cdots > X_1$ (Prop. 2.20). By definition of the Lex ordering, all monomials in f are only divisible by $X_1, \ldots, X_i$, so we may assume that f lies in I. This shows that every X_i is integral over $\mathbb{K}[X_{i+1}, \ldots, X_n]/I$. We get the requested result by induction on i: first, this is clear if $i = m$. Now assume that we know that $\mathbb{K}[X_i, \ldots, X_n]/I$ is an integral extension of $\mathbb{K}[X_{m+1}, \ldots, X_n]$. From the above, we also know that X_{i-1} is integral over $\mathbb{K}[X_i, \ldots, X_n]$, and so, since the composition of integral homomorphisms is integral, we get the requested result.

Finally, we want to check the second part of the definition of Noether position. Assume that there is a non-zero polynomial in $\mathbb{K}[X_{m+1}, \ldots, X_n] \cap I$. If this polynomial is constant, then I is not a proper ideal. And if this polynomial has degree at least 1, it is a non-trivial syzygy between $X_{m+1}, \ldots, X_n$, contradicting the regularity hypothesis.

(2 $\implies$ 4). For any $i \in \{1, \ldots, m\}$, write $f_i' = f_i(X_1, \ldots, X_m, 0, \ldots, 0)$. Since any permutation of a regular sequence is a regular sequence, $(X_{m+1}, \ldots, X_n, f_1, \ldots, f_m)$ is a regular sequence, that is, for any $1 \leq i \leq m$, f_i is not a zero divisor in

$$\mathbb{K}[X_1, \ldots, X_n]/\langle X_{m+1}, \ldots, X_n, f_1, \ldots, f_{i-1} \rangle$$

As a consequence, factoring in the quotient by $\langle X_{m+1}, \ldots, X_m \rangle$, f_i' is no zero-divisor in

$$\mathbb{K}[X_1, \ldots, X_m]/\langle f_1', \ldots, f_{i-1}' \rangle.$$

(4 $\implies$ 2). For any i, write $f_i = f_i' + r_i$ with $f_i' \in \mathbb{K}[X_1, \ldots, X_m]$, and $r_i \in \langle X_{m+1}, \ldots, X_n \rangle$. Let $1 \leq i \leq n$. Assume that $gf_i \in \langle X_{m+1}, \ldots, X_n, f_1, \ldots, f_{i-1} \rangle$:

$$\begin{aligned} gf_i = gf_i' + gr_i &= \sum_{j=1}^{i-1} g_j f_j + \sum_{j=m+1}^{n} g_j X_j \\ &= \sum_{j=1}^{i-1} g_j f_j' + R \qquad \text{with } R \in \langle X_{m+1}, \ldots, X_n \rangle. \end{aligned}$$

As a consequence, considering only the monomials in $\mathbb{K}[X_1, \dots, X_m]$

$$g' f'_i = \sum_{j=1}^{i-1} g_j f'_j \text{ where } g' = g(X_1, \dots, X_m, 0, \dots, 0).$$

Since $\theta_m(F)$ is regular, $g' \in \langle f'_1, \dots, f'_{i-1} \rangle$:

$$g = g' + r \in \langle f'_1, \dots, f'_{i-1} \rangle + \langle X_{m+1}, \dots, X_m \rangle = \langle f_1, \dots, f_{i-1} \rangle + \langle X_{m+1}, \dots, X_m \rangle.$$

And indeed, f_i is no zero-divisor in $\mathbb{K}[X_1, \dots, X_n]/\langle X_{m+1}, \dots, X_n, f_1, \dots, f_{i-1} \rangle$. It means that $(X_{m+1}, \dots, X_n, f_1, \dots, f_m)$ is a regular sequence. By permutation, we conclude that the sequence $(f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ is regular.

(4 $\iff$ 3). The sequence $\theta_m(F) = (\theta_m(f_1), \dots, \theta_m(f_m)) \in \mathbb{K}[X_1, \dots, X_m]^m$ is regular if and only if the sequence $(\theta_m(f_1), \dots, \theta_m(f_m), X_{m+1}, \dots, X_n)$ is regular. The equivalence between 3 and 4 is then a mirror of the equivalence between 1 and 2. □

1.2.4. Semi-regular sequences

The length of a regular sequence cannot be more than the number of indeterminates in the polynomial algebra. In applications, it is often required to solve so-called *overdetermined systems*, that is systems with more equations than unknowns. In the case of homogeneous systems, we use the notion of *semi-regularity* to characterize systems which are "as regular as they can".

Definition 1.45 (Semi-regular system)**.** Let m be an integer, and let $F = (f_1, \dots, f_m) \in A^m$ be a sequence of homogeneous polynomials. For any $i \in \mathbb{N}$, let $F_i = (f_1, \dots, f_i)$. We say that the sequence F is *semi-regular* if and only if for any $i \in \{1, \dots, m\}$ and any $d \in \mathbb{N}$, the linear map given by the multiplication by f_i:

$$\begin{array}{rccc} s_{i,d} : & (A/\langle F_{i-1} \rangle)_d & \longrightarrow & (A/\langle F_{i-1} \rangle)_{d+d_i} \\ & g & \longmapsto & f_i g \end{array}$$

is full-rank (either injective or surjective).

Proposition 1.46 ([Par10, Prop. 1])**.** *Let m be an integer, and let $F = (f_1, \dots, f_m) \in A^m$ be a sequence of homogeneous polynomials. For any $i \in \mathbb{N}$, let $F_i = (f_1, \dots, f_i)$.*

The sequence F is semi-regular if and only if for all $i \in \{1, \dots, m\}$, the Hilbert series of $A/\langle F_i \rangle$ is

$$\mathrm{HS}_{A/\langle F_i \rangle}(T) = \left\lfloor \frac{\prod_{i=1}^{m}(1 - T^{d_i})}{(1-T)^n} \right\rfloor$$

where $\lfloor \bullet \rfloor$ is the truncation before the first negative coefficient.

In particular, regular sequences are semi-regular, and if F is semi-regular, then all subsequences F_i are semi-regular. On the other hand, with this definition, permutations of a semi-regular sequence need not be semi-regular.

1.3. Algebraic geometry

Throughout this section, $\mathbb{K}$ is a field, $n \in \mathbb{N}$, and A is the polynomial algebra $\mathbb{K}[X_1, \ldots, X_n]$.

1.3.1. Nullstellensatz

Algebraic geometry establishes a correspondence between algebraic objects (ideals of polynomial algebras) and geometric objects (solution sets of polynomial equations). The backbone of this correspondence is Hilbert's Nullstellensatz (or theorem of zeroes).

Definition 1.47. Let I be an ideal of A, let $\mathbb{L}$ be a field containing $\mathbb{K}$. The *affine variety* (or in short *variety*) defined by I in $\mathbb{L}$ is the set

$$V_{\mathbb{L}}(I) = \{(\mathbf{x}) \in \mathbb{L}^n \mid \forall f \in I, f(\mathbf{x}) = 0\}.$$

When the field $\mathbb{L}$ is the algebraic closure of the field of definition $\mathbb{K}$, we simply write this set $V(I)$.

Let V be a subset of $\mathbb{L}^n$, we define the ideal associated with V as

$$I(V) = \{f \in \mathbb{L}[\mathbf{X}] \mid \forall (\mathbf{x}) \in V, f(\mathbf{x}) = 0\}$$

Theorem 1.48 (Hilbert's Nullstellensatz, [CLO07, Ch. 4, sec. 1, th. 2 and sec. 2, th. 4])**.** *Let $\mathbb{K}$ be an algebraically closed field. Then*

$$I(V(I)) = \sqrt{I}$$

where $\sqrt{I}$ is the radical of I.

1.3.2. Zariski topology

Algebraic varieties satisfy all axioms of the closed sets of a topology:

Proposition 1.49 ([Har77, Prop. 1.2])**.**

- $V_{\mathbb{L}}(0) = \mathbb{L}^n$
- $V_{\mathbb{L}}(A) = \varnothing$
- $V_{\mathbb{L}}(I) \cup V_{\mathbb{L}}(J) = V(I \cap J)$
- $\bigcap V_{\mathbb{L}}(I_i) = V(\sum I_i)$

Definition 1.50 (Zariski topology)**.** Taking all subsets $V_{\mathbb{L}}(I) \subset \mathbb{L}^n$ as closed sets, one obtains a topology on $\mathbb{L}^n$, called the *Zariski topology*.

Definition 1.51 (Irreducible subset)**.** A topological space is called *irreducible* if it cannot be written as the union of two proper closed subsets.

Equivalently, it means that two non-empty open subsets have non-empty intersection: in other words, all non-empty open subsets are dense.

Proposition 1.52 ([Har77, Cor. 1.4]). *If $\mathbb{L}$ is algebraically closed, irreducible varieties correspond to prime ideals of $\mathbb{L}[\mathbf{X}]$.*

Definition 1.53 (Irreducible component). Let $V \subset \mathbb{L}^n$ be a variety. There exists, up to permutation, a unique decomposition

$$V = V_1 \cup \cdots \cup V_k$$

such that all V_i are irreducible varieties and none contain another. The V_i are called the *irreducible components* of V.

Other topological notions (open set, closure, interior, boundary…) are defined as usual. *Locally closed* sets are defined as the intersection of an open set and a closed set. The last definition of this section is an interesting special case:

Definition 1.54 (Distinguished open subset). Let $V \subset \mathbb{L}^n$ be a variety and let $S \subset \mathbb{L}^n$ be a hypersurface, that is a variety defined by a single polynomial f. The open subset $V \setminus S$ is called a *distinguished open subset* of V.

1.3.3. Morphisms and rational maps

We shall be interested in functions between varieties, or between dense subsets of varieties. Because of the underlying algebraic structure, we can define polynomials on varieties, and the

Definition 1.55 (Morphism). Let $V_1 \subset \mathbb{K}^n$ and $V_2 \subset \mathbb{K}^m$ be varieties. A *morphism* $V_1 \longrightarrow V_2$ is an application φ which can be expressed, coordinate-wise, by polynomials:

$$\begin{array}{rccc} \varphi & : & V_1 & \longrightarrow & V_2 \\ & & (x_1, \ldots, x_n) & \longmapsto & (\varphi_1(\mathbf{x}), \ldots, \varphi_m(\mathbf{x})) \quad \text{with } \varphi_i \in \mathbb{K}[X_1, \ldots, X_n]/I(V_1) \end{array}$$

We say that V_1 and V_2 are *isomorphic* if there exist morphisms $V_1 \longrightarrow V_2$ and $V_2 \longrightarrow V_1$ which are mutually inverse.

Definition 1.56 (Rational map). Let $V_1 \subset \mathbb{K}^n$ and $V_2 \subset \mathbb{K}^m$ be varieties. A *rational map* $V_1 \dashrightarrow V_2$ is an application φ which can be expressed, coordinate-wise, by rational fractions:

$$\begin{array}{rccc} \varphi & : & U_1 & \dashrightarrow & V_2 \\ & & (x_1, \ldots, x_n) & \longmapsto & (\varphi_1(\mathbf{x}), \ldots, \varphi_m(\mathbf{x})) \quad \text{with } \varphi_i \in \mathbb{K}(X_1, \ldots, X_n) \end{array}$$

where U_1 is the Zariski-open subset of V_1 defined as the non-vanishing locus of the denominators of the φ_i.

Definition 1.57 (Birational equivalence). If V_1 and V_2 are *irreducible* varieties, we say that V_1 and V_2 are *birationally equivalent* (or simply *birational*) if there exist rational maps $V_1 \dashrightarrow V_2$ and $V_2 \dashrightarrow V_1$.

Recall that on irreducible varieties, non-empty open subsets are dense. So two irreducible varieties being birational means that, apart from some closed sets with empty interior, they are in a one-to-one rational correspondence.

A useful example of application of these definitions is the following classic construction, associating a distinguished open subset of a variety in $\mathbb{K}^n$ to an algebraic variety in $\mathbb{K}^{n+1}$:

Proposition 1.58. *Let $V \subset \overline{\mathbb{K}}^n$ be an irreducible variety, let $f \in \mathbb{K}[\mathbf{X}]$, and let W be the distinguished open subset $V \setminus V(f)$. Consider the variety $W' \subset \overline{\mathbb{K}}^{n+1}$, with coordinates $X_1, \ldots, X_n, U$ defined by $U \cdot f - 1 = 0$. Then there is an birational map with no pole:*

$$\begin{array}{ccc} W & \simeq & W' \\ (x_1, \ldots, x_n) & \longmapsto & \left(x_1, \ldots, x_n, \frac{1}{f(\mathbf{X})}\right) \\ (x_1, \ldots, x_n) & \longleftarrow & (x_1, \ldots, x_n, u) \end{array}$$

The ring $\mathbb{K}[\mathbf{X}, U]/\langle U \cdot f - 1\rangle$ is called the localization *of $\mathbb{K}[\mathbf{X}]$ at f, and denoted $\mathbb{K}[\mathbf{X}, f^{-1}]$ or $\mathbb{K}[\mathbf{X}]_f$.*

This is not true for general subsets, for example any rational map from $\mathbb{K}^2 \setminus \{(0,0)\}$ will necessarily have a curve of poles.

1.3.4. Dimension and degree

The dimension of an algebraic variety is defined as a topological dimension. We shall see that this definition coincides with the algebraic notion introduced earlier, and with the usual notion of dimension when working over the reals or the complex (in Section 1.3.7).

Definition 1.59 (Krull dimension, equidimensional varieties)**.** Let V be an irreducible variety of $\mathbb{L}^n$. The *Krull dimension* of V (or simply *dimension*) is the largest $d \in \{0, \ldots, n\}$ such that there exists a chain of irreducible closed subsets

$$V = V_1 \supsetneq V_2 \supsetneq \cdots \supsetneq V_d.$$

Let V be any variety, its dimension is defined as the maximal dimension of one of its irreducible components. Let X be any subset of $\mathbb{L}^n$, its dimension is defined as the dimension of its closure. As usual, the *codimension* of V is $n - d$.

Assume that $\mathbb{L}$ is algebraically closed. If all irreducible components of a variety V have the same dimension, we say that V is *equidimensional*. Any variety V is the union of a finite set of equidimensional varieties called *equidimensional components*: the d-equidimensional component of V is the union of all irreducible components of V having dimension d. The *local dimension* of V at a point $\mathbf{x} \in V$ is the largest dimension of an irreducible component of V containing $\mathbf{x}$.

Proposition 1.60. *If $\mathbb{L}_1 \subset \mathbb{L}_2$ are extensions of $\mathbb{K}$, then*

$$\dim(V_{\mathbb{L}_1}(I)) \leq \dim(V_{\mathbb{L}_2}(I))$$

Proof. Consequence of $V_{\mathbb{L}_1}(I) \subset V_{\mathbb{L}_2}(I)$. □

Proposition 1.61 ([Har77, Ex. 1.9])**.** *Let $f_1, \ldots, f_m \in \mathbb{K}[X_1, \ldots, X_n]$. All irreducible components of the variety $V_{\overline{\mathbb{K}}}$ have dimension at least $n - m$.*

Proposition 1.62 ([Har77, Prop. 1.7])**.** *Let I be an ideal with dimension d, then $V_{\overline{\mathbb{K}}}(I)$ has dimension d.*

In particular, if I is a 0-dimensional ideal, $V(I)$ is a finite set of points. The cardinality of this set is at most $\deg(I)$, and it is exactly $\deg(I)$ if points are counted with a notion of multiplicity and $\mathbb{K}$ is algebraically closed. For example, we shall see in Section 2.3.1 ideals generated by equations of the form

$$\begin{cases} g_1(X_1, \dots, X_n) & = X_1 + r_1(X_n) \\ g_2(X_2, \dots, X_n) & = X_2 + r_2(X_n) \\ & \vdots \\ g_{n-1}(X_{n-1}, X_n) & = X_{n-1} + r_{n-1}(X_n) \\ g_n(X_n) & = X_n^\delta + r_n(X_n) \end{cases}$$

For such an ideal, the quotient $\mathbb{K}[\mathbf{X}]/I$ is isomorphic to a univariate quotient:

$$\mathbb{K}[\mathbf{X}]/I \simeq \mathbb{K}[X_n]/\langle g_n \rangle$$

and the dimension of the right-hand side of the equality, as a $\mathbb{K}$-vector space, is the degree δ of the univariate polynomial g_n. So by definition, we have $\delta = \deg(I)$.

The system may not have exactly δ solutions in an algebraic closure of $\mathbb{K}$, but if it does not, it means that g_n has roots with multiplicity, and it makes sense to give the same multiplicity to the corresponding points in $V(I)$.

For positive dimensional systems, the degree is interpreted through Bézout's theorem. We shall give a rigorous statement of this theorem in Section 1.3.5, but informally-speaking, it means that the intersection of a variety of degree d_1 and a hypersurface of degree d_2 is expected to have degree $d_1 d_2$. Given a d-equidimensional variety $V(I)$ in $\mathbb{K}^n$, we may thus cut it with d random hyperplanes, in order to obtain a finite set of points whose cardinality (with multiplicities) is the degree $\deg(I)$.

1.3.5. Geometric interpretation of the regularity properties

In this section, we look back at the regularity properties defined in Section 1.2, in particular regular sequences and Noether position. We shall examine these properties through the algebra-geometry correspondence.

Proposition 1.63 ([Bar04, Prop. 1.7.4(1)], [Spa12, Prop. 1.43]). *Let m be an integer, W a system of weights, $F = (f_1, \dots, f_m) \in A^m$ a sequence of W-homogeneous polynomials, and $I = \langle F \rangle$. Then F is a regular sequence if and only if $V(I)$ is equidimensional with dimension $n - m$.*

Proof. From Prop. 1.61, we know that all irreducible components of $V(I)$ have dimension at least $n - m$, and from Prop. 1.62, with the Hilbert series described Prop. 1.38, $V(I)$ has dimension at most $n - m$, and thus all irreducible components of $V(I)$ as well.

Now assume that F is not regular. Let i be such that f_i is a zero divisor in $\mathbb{K}[\mathbf{X}]/\langle f_1, \dots, f_{i-1} \rangle$, and assume that i is minimal, that is $(f_1, \dots, f_{i-1})$ is a regular sequence. Let $I_{i-1} = \langle f_1, \dots, f_{i-1} \rangle$ and let $V_{i-1} = V(I_{i-1})$, this variety is equidimensional with dimension $n - i + 1$. Let I_f be the annihilator of f_i in A_{i-1}, that is

$$I_f = \{ g \in \mathbb{K}[\mathbf{X}]/I_{i-1} \mid g f_i = 0 \in \mathbb{K}[\mathbf{X}]/I_{i-1} \}.$$

Since f_i is a zero-divisor in $\mathbb{K}[\mathbf{X}]/I_{i-1}$, I_f is a non-zero ideal of $\mathbb{K}[\mathbf{X}]/I_{i-1}$ such that $fI_f = 0$ in $\mathbb{K}[\mathbf{X}]/I_{i-1}$. Geometrically, it means that

$$V_{i-1} = (V_{i-1} \cap V(f)) \cup (V_{i-1} \cap V(I_f)),$$

and since I_f is non-zero, $V(I_f) \cap V_{i-1}$ is a proper subvariety of V_{i-1}. Hence $V(f) \cap V_{i-1}$ is non-empty, and f vanishes on an irreducible component V of V_{i-1}. So $V(\langle f_1, \ldots, f_i \rangle)$ contains V, and has dimension $n - i + 1$. So $V(F)$ contains

$$V \cap V(\langle f_{i+1}, \ldots, f_m \rangle)$$

and by Prop. 1.61, it has dimension at least $n - m + 1 > n - m$. □

In particular, if F is a regular sequence of length n, the variety $V(\langle f \rangle)$ has dimension 0, which means that it is a finite set of points. Systems defined by a zero-dimensional regular sequences, sometimes called *square systems* (by analogy with linear algebra, where systems with as many equations as unknowns are defined by a square matrix), play a particular role in resolution strategies.

Theorem 1.64 (Bézout, consequence of [Spa12, Th. 1.67 and 1.68]). *Let $F = (f_1, \ldots, f_m)$ be a system of W-homogeneous polynomials with total degree $d_1, \ldots, d_n$. If F is a regular sequence, then the degree of the ideal $\langle F \rangle$ is given by the Bézout bound:*

$$\deg(I) = \frac{\prod_{i=1}^{m} d_i}{\prod_{i=1}^{n} w_i}.$$

In particular, if $m = n$, then the dimension of $A/\langle F \rangle$ as a $\mathbb{K}$-vector space is either infinite or equal to $d_1 \cdots d_n / w_1 \cdots w_n$.

Proof. The first statement is a consequence of the formula for the Hilbert series of a regular sequence. For the second statement, from Props. 1.61 and 1.63, $V(\langle F \rangle)$ has dimension 0 if and only if F is a regular sequence. In that case, the result is a consequence of the first statement. If $V(\langle F \rangle)$ has positive dimension, the dimension of $A/\langle F \rangle$ is infinite. □

All this can be applied in a non-homogeneous setting, by working with the homogenized system F^h instead. However, in that case the Bézout bound is really a bound, not necessarily sharp. The reason for that is that F^h may have more solutions than F. These solutions are called *solutions at infinity* and correspond to solutions with $H = 0$ (where H is the homogenization indeterminate). For example, two affine lines in a plane (degree 1 hypersurfaces) may not intersect if they are parallel.

Equivalently, solutions at infinity are points where the highest degree components of polynomials in F simultaneously vanish. We shall encounter this behavior again when describing algorithms for polynomial system solving, in Section 2.4.3, and we will define regularity in the affine sense, in order to ensure that this phenomenon does not appear.

We conclude this section with the following geometrical interpretation of Noether position.

Proposition 1.65 ([Eis95, Cor. 9.3]). *Assume that $\mathbb{K}$ is algebraically closed, and let $I \subset \mathbb{K}[\mathbf{X}]$ be an ideal in Noether position with respect to $X_1, \ldots, X_m$. Then the projection*

$$\begin{array}{rccc} \pi : & \mathbb{K}^n & \longrightarrow & \mathbb{K}^{n-m} \\ & (x_1, \ldots, x_n) & \longmapsto & (x_{m+1}, \ldots, x_n) \end{array}$$

is surjective, and the fibers $\pi^{-1}(\mathbf{y})$ ($\mathbf{y} \in \mathbb{K}^{n-m}$) are finite.

1.3.6. Singularities

The study of singularities of algebraic varieties and morphisms is of first importance for classification problems. We shall see an example of such application in Section 1.3.7 (Theorem 1.82). For now, in this section, we give some definitions and useful theorems.

In the following section, we assume that $\mathbb{K}$ is algebraically closed.

Definition 1.66 (Jacobian matrix)**.** Let $G = (g_1, \ldots, g_m) : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ be a polynomial map, and let $\mathbf{V} = \{V_1, \ldots, V_k\} \subset \{X_1, \ldots, X_n\}$. The *Jacobian matrix* of G with respect to $\mathbf{V}$ is the matrix of the differential of G:

$$\mathrm{Jac}(G)_{\mathbf{V}} = \begin{bmatrix} \frac{\partial g_1}{\partial V_1} & \cdots & \frac{\partial g_1}{\partial V_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial V_1} & \cdots & \frac{\partial g_m}{\partial V_n} \end{bmatrix}.$$

The Jacobian matrix of G with respect to $\mathbf{X}$ is written $\mathrm{Jac}(G)$.

Definition 1.67 (Smoothness)**.** Let V be a variety in $\mathbb{K}^n$, and let $F = (f_1, \ldots, f_m)$ be a system of generators of $I(V)$. Let $\mathbf{x} \in V$, and let d be the local dimension of V at $\mathbf{x}$. The point $\mathbf{x}$ is a *regular point* of V if the Jacobian matrix $\mathrm{Jac}(F)(\mathbf{x})$ has rank $n - d$. Otherwise, $\mathbf{x}$ is a *singular point* of V.

The *regular locus* (*resp. singular locus*) of V is the set $\mathrm{reg}(V)$ (*resp.* $\mathrm{sing}(V)$) of all regular (*resp.* singular) points of V. If $\mathrm{sing}(V) = \varnothing$, we say that V is *smooth*.

Remark 1.68. This definition of regularity requires F to be a system of generators of $I(V)$. By the Nullstellensatz, it is equivalent to saying that $V = V(\langle F \rangle)$ and $\langle F \rangle$ is *radical*.

Definition 1.69 (Critical point)**.** With the same notations as above, let $g = (g_1, \ldots, g_k)$ define a polynomial map $\mathbb{K}^n \longrightarrow \mathbb{K}^k$. We consider the vertical joint matrix

$$\mathrm{Jac}(f, g) = \begin{bmatrix} \mathrm{Jac}(f) \\ \mathrm{Jac}(g) \end{bmatrix}.$$

Let $\mathbf{x} \in \mathrm{reg}(V)$. The point $\mathbf{x}$ is a *critical point* of the map g restricted to V if $\mathrm{Jac}(f, g)(\mathbf{x})$ has rank less than $k + n - d$. In that case, its image $g(\mathbf{x})$ is a *critical value* of g restricted to V. A point $\mathbf{y} \in \mathbb{K}^k$ which is not a critical value of g restricted to V is a *regular value* of g restricted to V.

The set of critical points of g restricted to V is denoted by $\mathrm{crit}(G, V)$.

Remark 1.70. If $\mathbf{x} \in \mathrm{sing}(V)$, and $\mathbf{x}$ lies in a d-dimensional irreducible component of V, the rank of $\mathrm{Jac}(f, g)(\mathbf{x})$ is necessarily less than $k + n - d$. In other words, if V is d-equidimensional, the locus at which $\mathrm{Jac}(f, g)(\mathbf{x})$ has rank less than $k + n - d$ is $\mathrm{sing}(V) \cup \mathrm{crit}(g, V)$.

Remark 1.71. Let $\mathbf{y} \in \mathbb{K}^k$; if $\mathbf{y}$ is a regular value of g restricted to V, then $\langle f, (g_i - y_i)\rangle$ is radical and $g^{-1}(\mathbf{y}) \cap V \subset \mathbb{K}^n$ is smooth. In particular, if $0 \in \mathbb{K}^k$ is a regular value of g restricted to V, then the variety $V(\langle f, g\rangle) \subset \mathbb{K}^n$ is smooth.

The Jacobian criterion is a very important tool linking the Jacobian matrix and properties of the underlying variety:

Theorem 1.72 (Jacobian criterion, [Nal15, Th. 1.9]). *Let $V \in \mathbb{K}^n$ be an algebraic variety, and $f = (f_1, \ldots, f_m) \subset \mathbb{K}[\mathbf{X}]$ a system of generators of $I(V)$.*

1. *If V is d-equidimensional, then for all $\mathbf{x} \in \mathbb{K}^n$, $\mathbf{x}$ is a regular point of V if and only if $\mathrm{Jac}(f)(\mathbf{x})$ has rank $n - d$;*
2. *If for all $\mathbf{x} \in \mathbb{K}^n$, $\mathrm{Jac}(f)(\mathbf{x})$ has rank $n - d$, then V is d-equidimensional and smooth.*

We conclude this section with two powerful tools when working with critical values.

Sard's theorem states that given a polynomial map, a generic point of the target space is a regular value.

Theorem 1.73 (Sard's theorem, [Nal15, Lemma 1.16]). *Let $V \subset \mathbb{K}^n$ be a variety, and let $g : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ be a polynomial map. Then $g(\mathrm{crit}(g, V))$ is contained in a proper hypersurface of $\mathbb{K}^m$.*

Thom's transversality theorem states that given a smooth variety, the intersection with generic coordinate level lines is smooth.

Theorem 1.74 (Thom's weak transversality theorem, [Nal15, Th. 1.18],[Demoo, Th. 3.7.4]). *Let $f : \mathbb{K}^n \times \mathbb{K}^k \longrightarrow \mathbb{K}^r$ be a polynomial map. For $\mathbf{y} \in \mathbb{K}^k$, define the partial application*

$$\begin{array}{rccc} f_{\mathbf{y}} : & \mathbb{K}^n & \longrightarrow & \mathbb{K}^r \\ & \mathbf{x} & \longmapsto & f(\mathbf{x}, \mathbf{y}) \end{array}$$

Let $\mathcal{U} \subset \mathbb{K}^n$ be an open subset such that $0 \in \mathbb{K}^r$ is a regular value of f restricted to $\mathcal{U} \times \mathbb{C}^k$. Then there exists an open subset $\mathcal{V} \subset \mathbb{K}^k$ such that for all $\mathbf{y} \in \mathcal{V}$, $0 \in \mathbb{K}^r$ is a regular value of $f_{\mathbf{y}}$ restricted to $\mathcal{U}$.

1.3.7. Real semi-algebraic geometry

In this section, we focus on the case where the base field is $\mathbb{R}$. The definitions of the previous sections allow us to work with real algebraic sets (defined as subsets of $\mathbb{R}^n$ defined as the sets of zeroes of some polynomial ideal).

However, geometry over the reals is richer if we take into account the fact that $\mathbb{R}$ is an ordered field, by allowing inequalities in the definition of the basic objects of the geometry: that is the definition of a real semi-algebraic set.

Definition 1.75. A *real semi-algebraic set* is a finite union V of subsets V_i, with each V_i defined by a set of polynomial equations $\{F_i = 0 \mid F_i \subset \mathbb{R}[\mathbf{X}]\}$ and a set of polynomial inequations $\{G_i > 0 \mid G_i \subset \mathbb{R}[\mathbf{X}]\}$, as

$$V_i = \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \forall\, f \in F_i, f(\mathbf{x}) = 0 \\ \forall\, g \in G_i, g(\mathbf{x}) > 0 \end{array} \right\}$$

Like algebraic varieties, real semi-algebraic sets are stable under finite union and intersection. Furthermore, they are also stable under complement. Finally, they are stable under projection onto linear subspaces:

Theorem 1.76 (Tarski-Seidenberg, [Nal15, Th. 1.13]). *Let V be a semi-algebraic subset of $\mathbb{R}^{n+1}$ and $\pi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n$ the projection onto the first n coordinates. Then $\pi(V)$ is a semi-algebraic subset of $\mathbb{R}^n$.*

Let V be a real semi-algebraic set of $\mathbb{R}^n$. Locally (up to intersection with an open subset of $\mathbb{R}^n$), V is a submanifold of $\mathbb{R}^n$, as defined by differential geometry:

Definition 1.77 (Submanifold). A submanifold M is a subset of $\mathbb{R}^n$ such that, for all $\mathbf{x} \in M$, there exists an open subset $U \subset \mathbb{R}^n$ such that $M \cap U$ is diffeomorphic to $\mathbb{R}^d$, for some integer $d \in \{0, \dots, n\}$. In this case, d is the *local dimension* of M at $\mathbf{x}$. The *dimension* of M is the largest local dimension at a point of M.

Remark 1.78. Let V be a real algebraic set. This definition of dimension coincides with the Krull dimension, as defined in Section 1.3.4.

Remark 1.79. One can define submanifolds of $\mathbb{C}^n$ in a similar way. Then complex algebraic varieties are locally submanifolds of $\mathbb{C}^n$. Moreover, the dimension of an algebraic variety V seen as a submanifold is the Krull dimension of V, and if V is d-equidimensional (for example irreducible), for all points at which V is a submanifold, V has local dimension d.

Remark 1.80. Over the reals, it is no longer true: *Whitney's umbrella* (Figure 1.2) is a classic example of an irreducible subset where points may have different local dimensions. It is the algebraic set W of $\mathbb{R}^3$ defined as the real zeroes of $X^2 - Y^2Z$. This polynomial is irreducible (over $\mathbb{C}$), so the algebraic variety W is irreducible.

However, W is the disjoint union of two semi-algebraic sets:

$$W_+ = \left\{ (x,y,z) \;\middle|\; \begin{array}{l} x^2 = y^2 z \\ z \geq 0 \end{array} \right\}$$
$$W_- = \left\{ (x,y,z) \;\middle|\; \begin{array}{l} x = y = 0 \\ z < 0 \end{array} \right\}$$

The semi-algebraic set W_+ has dimension 2 (the projection onto (x, y) is a local diffeomorphism), and the semi-algebraic set W_- has dimension 1 (it is a half-line).

The final result of this section is an algebraic version of Ehresmann's theorem, or Thom's isotopy lemma, giving a description of fibers of the projection of a real semi-algebraic set under some hypotheses.

We first recall the following topological definition:

Definition 1.81 (Proper map). Let A and B be two topological spaces, and $f : A \longrightarrow B$ a map. The map f is *proper* if and only if for all compact subsets $K \subset B$, $f^{-1}(K)$ is compact in A.

Theorem 1.82 (Thom's isotopy lemma, [CS95]). *Let V be a real semi-algebraic set of $\mathbb{R}^n \times \mathbb{R}^t$, and let $\pi : \mathbb{R}^n \times \mathbb{R}^t \longrightarrow \mathbb{R}^t$ be the projection onto the last t coordinates. Assume that:*

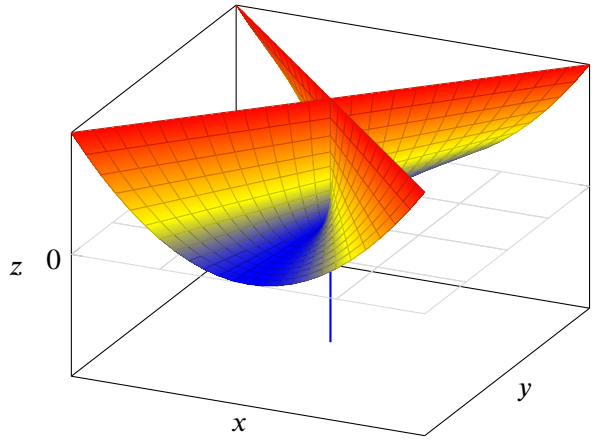

Figure 1.2: Whitney's umbrella

- *V is smooth;*
- *V is locally closed;*
- *the restriction $\pi_{|V}$ has no critical points;*
- *$\pi_{|V}$ is proper (for the euclidean topology);*
- *V is t-equidimensional.*

Then there exists a finite set F and a semi-algebraic diffeomorphism

$$h = (h_0, \pi) : V \xrightarrow{\sim} F \times \mathbb{R}^t.$$

1.4. Genericity

1.4.1. Definition

Genericity formalizes the intuitive notion that some property is *almost always* true.

Definition 1.83. Let m be an integer, and $S = (S_i)_{i\in\{1,\dots,m\}}$ a family of finite subsets of monomials of A. Let $A_{m,S}$ be the set of families of m polynomials $f_i \in \mathbb{K}[\mathbf{X}]$, with $\mathrm{Supp}(f_i) \subset S_i$ for all $i \in \{1,\dots,m\}$. Elements of $A_{k,S}$ can be seen as points in an affine space over $\overline{\mathbb{K}}$, the coordinates being the coefficients of the polynomials. Finally, let P be a property of elements of $A_{m,S}$.

We say that P is *generic* amongst elements of $A_{m,S}$ if the set of families satisfying P in $A_{m,S}$ contains a non-empty Zariski-open subset of $A_{m,S}$.

A typical choice for the support S_i is the set of polynomials with degree less than a given integer d, or homogeneous polynomials with degree exactly d.

1.4.2. Generic properties of homogeneous systems

We will now prove genericity results for the regularity properties defined in Section 1.1.5, for homogeneous systems (with the total degree). More generally, the proofs that these properties cover Zariski-open subsets hold for any gradation, but for non-trivial systems of weights, these subsets need not be non-empty (see Section 3.2.4).

We assume that A is graded in total degree. For $d \in \mathbb{N}$, let A_d be the set of homogeneous generic polynomials with degree exactly d.

Proposition 1.84 ([Bar04, Prop. 1.7.4(5)], [Par10]). *Let $m \in \{1, \dots, n\}$ and let $(d_1, \dots, d_m) \in \mathbb{N}^m$ be a family of degrees. Homogeneous regular sequences with degree $(d_1, \dots, d_m)$ are generic amongst families of homogeneous polynomials of degree $(d_1, \dots, d_m)$, that is in $\prod_{i=1}^m A_{d_i}$.*

Proof. Assume that $m = n$ and let $F = (f_1, \dots, f_m) \in \mathbb{K}[\mathbf{A}][\mathbf{X}]$ be a sequence of polynomials in $\mathbf{X}$, with generic coefficients $\mathbf{A}$. Let $I = \langle F \rangle$. Write

$$H_{\text{reg}} = \frac{\prod_{i=1}^n (1 - T^{d_i})}{(1-T)^n} = \sum_{i=1}^{\infty} a_i T^i \text{ with } a_{\delta+1} = a_{\delta+2} = \cdots = 0$$

and let b_i be the coefficient of degree i in the Hilbert series of F.

By Prop. 1.38, for all $i \in \mathbb{N}$, $b_i \geq a_i$ and F is regular if and only if they are equal. For each i, this equality means that the rank of the Macaulay matrix of F at degree i is a_i, and this can be encoded by the non-vanishing of the $a_i \times a_i$-minors of this matrix. Let U_i be the Zariski-open subset defined by the non-vanishing of at least one of these minors.

Furthermore, if $b_{\delta+1} = 0$, it means that I contains all monomials of degree $\delta + 1$, and it implies that $b_{\delta+2} = \cdots = 0$. So we only need to check a finite number of equalities $a_i = b_i$, for i ranging from 0 to $\delta + 1$. Geometrically, it means that the set of regular sequences is the finite intersection U of the open sets U_i.

Consider the sequence $(X_1^{d_1}, \dots, X_n^{d_n})$, it is regular, and so U is non-empty.

If $m < n$, F is regular if and only if there exist linear forms $l_{m+1}, \dots, l_n$ such that the sequence $(f_1, \dots, f_m, l_{m+1}, \dots, l_n)$ is regular. In particular, the set of regular sequences of homogeneous polynomials with degree $(d_1, \dots, d_m)$ contains the set of sequences $F = (f_1, \dots, f_m)$ such that $(f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ is regular. From the above, regular sequences with degree $(d_1, \dots, d_m, 1, \dots, 1)$ form a non-empty Zariski-open subset U of $\prod_{i=1}^m A_{d_i} \times A_1^{n-m}$, and the closed subset C defined by $f_{m+1} = X_{m+1}, \dots, f_n = X_n$ does meet U. Consider the projection onto the m first components

$$\pi : \prod_{i=1}^m A_{d_i} \times A_1^{n-m} \longrightarrow \prod_{i=1}^m A_{d_i}$$

It is surjective, and its restriction to C is an isomorphism. Since $C \cap U$ is non-empty, $\pi(C \cap U)$ is a non-empty open subset of $\prod_{i=1}^m A_{d_i}$, defining exactly those sequences $F = (f_1, \dots, f_m)$ such that $(f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ is regular. □

Corollary 1.85. *Homogeneous sequences in (simultaneous) Noether position with respective degree $(d_1, \dots, d_m)$ are generic amongst families of homogeneous polynomials of degree $(d_1, \dots, d_m)$.*

Proof. This is a direct consequence of the proof of the case $m < n$ of Prop. 1.84. □

Corollary 1.86. *Let m be an integer, and let $(d_1, \ldots, d_m)$ be a family of degrees. Assume that there exists at least one semi-regular sequence with degree $(d_1, \ldots, d_m)$, then homogeneous semi-regular sequences with degree $(d_1, \ldots, d_m)$ are generic amongst families of homogeneous polynomials of degree $(d_1, \ldots, d_m)$.*

Proof. If $m \leq n$, semi-regular sequences are actually regular, and so Prop. 1.84 applies. Otherwise, semi-regular sequences define zero-dimensional ideals and are characterized by their Hilbert series, so the proof that zero-dimensional regular sequences form a Zariski-open subset in Prop. 1.84 can be repeated for semi-regular sequences. □

This corollary states that semi-regular sequences form a Zariski-open subset of all sequences, and that the property is generic if and only if this Zariski-open subset is non-empty. This is an open question:

Conjecture 1.87 (Fröberg,[Frö85])**.** Let m be an integer, and let $(d_1, \ldots, d_m)$ be a family of degrees. Then there exists at least one semi-regular sequence of homogeneous polynomials with respective degrees $d_1, \ldots, d_m$.

This conjecture is only proved in a handful of cases [Moro3, Th. 1.5]:

- $m \leq n$: then semi-regular sequences are regular sequences;
- $n = 2$;
- $n = 3$ and the field $\mathbb{K}$ is infinite;
- $m = n + 1$ and the field $\mathbb{K}$ has characteristic 0.

1.4.3. Generic changes of coordinates

Some properties are also true *up to a generic change of coordinates*: given a property P, this means that given any system F, the set of linear changes of coordinates L such that $F \circ L$ satisfies P is generic amongst all linear changes of coordinates.

A classic result of this kind is Noether's normalization lemma: it is a fundamental result of dimension theory, stating that given a variety V of dimension d, up to a generic change of coordinates, the projection $V \longrightarrow \mathbb{K}^{n-m}$ is surjective and finite.

We give here an algebraic transcription of this theorem for a regular sequence:

Theorem 1.88 (Noether's normalization theorem, [Eis95, Th. 13.3])**.** *Let $\mathbb{K}$ be an infinite field, F a regular sequence in $\mathbb{K}[X_1, \ldots, X_n]$, and $k \in \{1, \ldots, n\}$. Then, for a generic choice of $n - k$ linear forms $l_i(X_{i+1}, \ldots, X_n)$, $i \in \{1, \ldots, k\}$, the change of variables*

$$X_i = X'_i + l_i(X_{i+1}, \ldots, X_n)$$

is such that $F(X_1(\mathbf{X}'), \ldots, X_k(\mathbf{X}'), X_{k+1}, \ldots, X_n)$ is in Noether position with respect to the variables $X_{k+1}, \ldots, X_n$.

More generally, if F has length n, then for a generic choice of n linear forms $l_i(X_{i+1}, \dots, X_n)$, $i \in \{1, \dots, n\}$, the change of variables

$$X_i = X'_i + l_i(X_{i+1}, \dots, X_n)$$

is such that $F(X_1(\mathbf{X}'), \dots, X_n(\mathbf{X}'))$ is in simultaneous Noether position with respect to the order $X'_1 > \cdots > X'_n$.

Noether's normalization theorem is really made of two parts: first, up to a substitution $X_i = P(\mathbf{X})$, any system can be put in Noether position; and second, that under some hypotheses, we can make additional assumptions on the substitution. Here, the second part is that if $\mathbb{K}$ is infinite, then any generic linear change of variables gives a suitable transformation; this part of the proof of [Eis95, Th. 13.3] relies on the following lemma:

Lemma 1.89 (Noether's normalization lemma, [Eis95, lem. 13.2.c])**.** *Let $\mathbb{K}$ be an infinite field and $f \in A = \mathbb{K}[X_1, \dots, X_r]$ a non-constant polynomial. Then there are homogeneous elements $X'_1, \dots, X'_{r-1} \in A$ with degree 1 such that A is a finitely generated module over $\mathbb{K}[X'_1, \dots, X'_{r-1}, f]$. Furthermore, there exists a dense Zariski-open subset $U \subset \mathbb{K}^{r-1}$ such that for all $(a_i) \in U$, one can choose $X'_i = X_i - a_i X_r$.*

Proof. For any $1 \leq i \leq r-1$, let $a_i \in \mathbb{K}$, and let $X'_i = X_i - a_i X_r$. If f has degree d, let f_d be the degree d component of f. We need to show that for generic a_i, under this change of variables, f is monic in X_r, and it is enough to prove it for f_d:

$$\begin{aligned} f_d(X_1, \dots, X_r) &= f_d(X'_1 + a_1 X_r, X'_2 + a_2 X_r, \dots, X_{r-1} + a_{r-1} X_r) \\ &= f_d(a_1, \dots, a_{r-1}, 1) X_r^d + \dots \end{aligned}$$

So the set of all a_i's such that f is monic in X_r is exactly the set of all a_i's such that

$$f_d(a_1, \dots, a_{r-1}, 1) \neq 0,$$

and since f is homogeneous non-constant, this is a non-empty open subset of $\mathbb{K}^{r-1}$. □

If F is homogeneous, it is still homogeneous after any linear change of variables, so Noether's normalization theorem states that a regular homogeneous sequence can be made into a homogeneous sequence in Noether position, up to a generic change of variables. We shall identify suitable changes of variables for the same property to hold in a weighted setting in Section 3.2.3.

1.5. Determinantal varieties

In this section, we will define determinantal varieties, that is varieties defined as the locus of rank defects of a matrix with polynomial entries. These systems arise in many applications, for example in optimization, when using the Jacobian criterion to model singularities and critical points in terms of the minors of the Jacobian matrix.

1.5.1. Definition

Definition 1.90 (Determinantal ideal, determinantal variety)**.** Let r, k, n be integers such that $r \leq k$. Let $M \in \mathbb{K}[X_1, \ldots, X_n]^{k \times k}$ be a matrix with polynomial coefficients. The r'th *determinantal ideal* associated with M is the ideal $\mathcal{D}_r$ generated by its $(r+1)$-minors. The corresponding *determinantal variety* is

$$V_r = \{\mathbf{x} \in \mathbb{K}^n \mid \operatorname{rank}(M(\mathbf{x})) \leq r\}$$

In Chapter 4, we shall be interested in the study of the singularities of these varieties. Because of the structure of determinantal ideals, these singularities are strongly related to other determinantal varieties.

Proposition 1.91. *If $\mathcal{D}_r$ is radical, then*

$$V_{r-1} \subset \operatorname{sing}(V_r)$$

Proof. We will prove the following stronger statement: let $\mathbf{x} \in V_{r-1}$, then all partial derivatives of all $(r+1)$-minors of M vanish at $\mathbf{x}$. This will prove that the Jacobian of V_r at $\mathbf{x}$ is the zero matrix, and in particular has rank 0. Then, by hypothesis the ideal of all $(r+1)$-minors of M is radical, so we can use the Jacobian criterion to characterize $\operatorname{sing}(V)$, and we can conclude that $\mathbf{x} \in \operatorname{sing}(V_r)$.

To prove this statement, let $\mathbf{x} \in V_{r-1}$. For any $(r+1) \times (r+1)$-submatrix of M, the result we want to prove only depends on the coefficients of the submatrix. So *w.l.o.g.*, we may assume that $r = k-1$, let $D = \det(M)$. Let $\mathbf{x} \in V_{k-2}$, this means that all $(k-1)$-minors of M vanish at $\mathbf{x}$. Consider a matrix of polynomial indeterminates $\mathbf{U}$:

$$\tilde{M} = \begin{bmatrix} U_{1,1} & \ldots & U_{1,k} \\ \vdots & & \vdots \\ U_{k,1} & \ldots & U_{k,k} \end{bmatrix} \tag{1.2}$$

and let $\tilde{D}$ be its determinant. Then for any $i, j \in \{1, \ldots, k\}$,

$$\frac{\partial \tilde{D}}{\partial U_{i,j}} = (-1)^{i+j} \cdot \tilde{\mathcal{M}}_{i,j}(\mathbf{U})$$

where $\tilde{\mathcal{M}}_{i,j}$ is the $(k-1)$-minor of $\tilde{M}$ obtained by removing row i and column j.

For any $i, j \in \{1, \ldots, k\}$, let $m_{i,j}$ (*resp.* $\mathcal{M}_{i,j}$) be the coefficient at row i and column j of the matrix M (*resp.* the minor obtained by removing row i and column j from M). By the derivation chain rule, for any $X \in \{X_1, \ldots, X_n\}$,

$$\begin{aligned} \frac{\partial D}{\partial X} &= \sum_{i,j=1}^{k} (-1)^{i+j} \frac{\partial m_{i,j}}{\partial X} \cdot \tilde{\mathcal{M}}_{i,j}(\mathbf{m}) \\ &= \sum_{i,j=1}^{k} (-1)^{i+j} \frac{\partial m_{i,j}}{\partial X} \cdot \mathcal{M}_{i,j}(\mathbf{X}) \end{aligned}$$

Since by hypothesis all $(k-1)$ minors of M vanish at $\mathbf{x}$, all partial derivatives $\frac{\partial \Delta}{\partial X}$ vanish at $\mathbf{x}$. □

Generically, this inclusion is an equality.

Proposition 1.92 ([HNS15b, Prop. 2]). *Let $d \in \mathbb{N}$. Generically amongst matrices with coefficients of degree at most d:*

1. *$V_{r-1} = \mathrm{sing}(V_r)$;*
2. *the variety V_r is equidimensional with codimension $c = (k-r)^2$ if $n \geq (k-r)^2$, and empty otherwise;*
3. *the ideal D_r is radical.*

Proof. Let $\tilde{D}_r$ be the r'th determinantal ideal associated to the matrix $\tilde{M}$ as defined in Equation (1.2), and, as before, let $m_{i,j}$ be the coefficient at row i and column j in M. In $\mathbb{K}[\mathbf{X}, \mathbf{U}]$, the two determinantal ideals are related as

$$\langle D_r \rangle = \langle \tilde{D}_r \rangle + \langle U_{i,j} - m_{i,j}(\mathbf{X}) \rangle_{1 \leq i,j \leq k}.$$

Now, all three statements of the proposition are unconditionally true for $\tilde{M}$: see [BV88, Th. 2.10] for the radicality of $\tilde{D}_r$, and then [BV88, Prop. 1.1] for the other two statements. We shall prove that generically, these statements transfer to M.

Consider the application

$$\begin{array}{rccc} \mathcal{M}_{\mathbf{m}} : & \mathbb{K}^n & \longrightarrow & \mathbb{K}^{k^2} \\ & \mathbf{x} & \longmapsto & (m_{i,j}(\mathbf{x}))_{i,j} \end{array}$$

By Thom's weak transversality lemma, generically over the space of polynomials $(m_{i,j})$,

$$V_r \smallsetminus V_{r-1} = \mathcal{M}_{\mathbf{m}}^{-1}(\tilde{V}_r \smallsetminus \tilde{V_{r-1}})$$

is either empty or smooth and equidimensional with codimension $c = (k-r)^2$. This proves statements 1 and 2. It also proves that $V_r = V(D_r)$ satisfies the hypotheses of [Eis95, Th. 18.18].

Furthermore, add the substitution equations $U_{i,j} - m_{i,j}(\mathbf{X})$ to the generators of $\tilde{D}_r$ in $\mathbb{K}[\mathbf{U}, \mathbf{X}]$, and consider the Jacobian of this system. It writes as a block matrix

$$\mathrm{Jac}_{\mathbf{U},\mathbf{X}}(D_r) = \begin{bmatrix} \mathrm{Jac}_{\mathbf{U}}(\tilde{D}_r) & 0 \\ \mathrm{Id}_{k^2} & \mathrm{Jac}_{\mathbf{X}}(m_{i,j}) \end{bmatrix}$$

where the top-left block has $\binom{k}{r}$ rows and k^2 columns, and the bottom-right block has k^2 rows and n columns. By Sard's lemma, the set of $\mathbf{x}$ such that $\mathrm{Jac}_{\mathbf{X}}(m_{i,j})$ has non-full rank has codimension at least 1 in $\mathbb{K}^n$; and from the above, outside of $\tilde{V_{r-1}}$, $\mathrm{Jac}_{\mathbf{U}}(\tilde{D}_r)$ has rank c. So by Serre's Criterion for radicality ([Eis95, Th. 18.15(a) with Th. 18.18]), under the same genericity condition as above, the ideal D_r is radical, which completes the proof. □

1.5.2. Cramer's formula and Schur's complement

In particular, results from the previous section imply that the system of all $(r+1)$-minors of M is not a regular sequence (it has length $\binom{k}{r+1}^2$ and, generically, defines a variety with codimension $(k-r)^2$). However, it is possible to cover determinantal varieties with dense open subsets, each being an open subset of a complete intersection.

These constructions stem from the linear algebraic interpretation of the rank: a $k \times k$ matrix has rank at most r if and only if its kernel has dimension at least $k - r$. Algebraically, the correspondence between linear algebra and minors comes from Cramer's formula, and its generalization Shur's complement. Cramer's formula is a classical formula for solving invertible linear system. Here, we give an equivalent statement for matrices with corank 1.

Proposition 1.93 (Cramer's formula)**.** *Let R be a ring, $k \in \mathbb{N}$, $M \in R^{k\times k}$ an invertible matrix, $b \in R^k$ a vector, and consider the linear system in $x = {}^{\mathsf{T}}(x_1, \dots, x_k)$:*

$$Mx = b$$

Then this system has a unique solution, defined by

$$\forall\, i \in \{1, \dots, k\}, x_i = \frac{\det(M_i)}{\det(M)}$$

where M_i is the matrix obtained by replacing the column i in M with b.

Amongst its consequences, we recall the classic formula for the inverse of a matrix:

Proposition 1.94. *With the same notations and hypotheses,*

$$M^{-1} = \frac{1}{\det(M)}{}^{\mathsf{T}}\operatorname{Com}(M) \tag{1.3}$$

where $\operatorname{Com}(M)$ *is the* cofactor matrix *of M, whose coefficient at row i and column j is*

$$c_{i,j} = (-1)^{i+j} \cdot \ (k-1)\text{-minor of } M \text{ obtained by removing row } i \text{ and column } j.$$

Conversely, if M is such that $\det(M) \neq 0$*, then M is invertible in the localized ring $R_{\det(M)} = R[1/\det(M)]$, and its inverse is given by the same formula.*

Cramer's rule can be used to parameterize the kernel of non-invertible matrices. For example, assume that M is a $k \times k$ singular matrix such that the submatrix M_0 obtained by removing row k and column k is invertible, so that M has rank $k-1$. Let $x = {}^{\mathsf{T}}(x_1, \dots, x_k) \in \operatorname{Ker}(M)$:

$$M \cdot x = 0$$

Keeping only the $k-1$ first rows of M, the system becomes

$$M_0 \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_{k-1} \end{bmatrix} = -x_k C_k$$

where C_k is the k'th column in M. Cramer's rule implies that the kernel of M is parameterized by x_k as

$$\forall\, i \in \{1, \ldots, k-1\}, x_i = \frac{-x_k \det(M_i)}{\det(M_0)}$$

where M_i ($i \in \{1, \ldots, k-1\}$) is the submatrix obtained by replacing column i in M_0 with the last column C_k of M, less row i.

This construction is what we use to fill the gap between determinantal ideals and kernels of matrices. For matrices of smaller rank, it can be generalized:

Definition 1.95 (Schur complement)**.** Let $M \in R^{k\times k}$ be a square matrix, $p \in \{1, \ldots, k\}$, and A, B, C and D be respectively $p \times p$, $p \times k-p$, $k-p \times p$ and $k-p \times k-p$ matrices such that

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

Further assume that A is invertible. Then the *Schur complement* of the block A in M is the $k-p \times k-p$ matrix

$$M/A = D - CA^{-1}B. \tag{1.4}$$

Proposition 1.96. *The Schur complement of A in M appears in the following block diagonal decomposition of M:*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \mathrm{Id}_p & 0 \\ CA^{-1} & \mathrm{Id}_{k-p} \end{bmatrix} \cdot \begin{bmatrix} A & 0 \\ 0 & M/A \end{bmatrix} \cdot \begin{bmatrix} \mathrm{Id}_p & A^{-1}B \\ 0 & \mathrm{Id}_{k-p} \end{bmatrix}$$

Proposition 1.97. *The entry at row i and column j in M/A is $\mathcal{M}_{p,i,j}/\det(A)$, where $\mathcal{M}_{p,i,j}$ is the $(p+1)$-minor of M obtained by taking rows $1, 2, \ldots, p, p+i$ and columns $1, 2, \ldots, p, p+j$, or equivalently, the $(p+1)$-minor of M obtained by appending row i and column j to A.*

Proof. Let $a_{i,j}$, $a'_{i,j}$, $b_{i,j}$, $c_{i,j}$, $d_{i,j}$, $d'_{i,j}$ be the coefficient at row i and column j in respectively A, A^{-1}, B, C, D, M/A. Formula (1.4) yields that

$$d'_{i,j} = d_{i,j} - \sum_{u=1}^{p} c_{i,u} \sum_{v=1}^{p} a'_{u,v} b_{v,j}$$

By Formula (1.3),

$$a'_{u,v} = \frac{(-1)^{u+v} \mathcal{A}_{v,u}}{\det(A)}$$

where $\mathcal{A}_{v,u}$ is the $(p-1)$-minor of A obtained by removing row v and column u (u and v are swapped because of the transposition in Formula (1.3)). So, by decomposition over column j, $\det(A) \sum_{v=1}^{p} a'_{u,v} b_{v,j}$ is the p-minor of M obtained by removing column u in A and appending column j in B, and by decomposition over row i,

$$\det(A) d'_{i,j} = \det(A) d_{i,j} - \det(A) \sum_{u=1}^{p} c_{i,u} \sum_{v=1}^{p} a'_{u,v} b_{v,j}$$

is the $(p+1)$-minor of M obtained by adding row $p+i$ and column $p+j$ in M to A. □

To conclude this section, we give an immediate consequence of these last two results, which is a first local description of determinantal systems.

Proposition 1.98. *Assume that V_{r-1} has empty interior in V_r. Let A be a $r \times r$ submatrix of M. Then the open subset W of V_r at which $\det(A)$ does not vanish is defined by the vanishing of all $(r+1)$-minors of M containing A, and generically, they form a regular sequence.*

Proof. Up to permutation of the rows and columns, we may assume that A is the top-left $r \times r$ submatrix of M. If A is invertible, M has rank at least r, and by Proposition 1.96, M has rank r if and only if M/A is 0. By Proposition 1.97, the entries of this matrix are precisely the $(r+1)$ minors containing A.

Since generically, the variety V_r is equidimensional with dimension $(k-r)^2$, the same holds for the non-empty open subset W. The sequence of $(r+1)$-minors containing A has length $(k-r)^2$, so by the characterization of Propostion 1.63, they form a regular sequence. □

1.5.3. Incidence varieties

Incidence varieties offer an alternative modelling of determinantal varieties.

Definition 1.99. Let $r \in \{0, \dots, k-1\}$. The *incidence variety of rank r* associated with M is the variety $\mathcal{V}_r \subset \mathbb{K}^n \times (\mathbb{P}^{k-1}(\mathbb{K}))^{k-r}$ defined by:

$$M \cdot \begin{bmatrix} Y_{1,1} & \dots & Y_{1,k-r} \\ \vdots & & \vdots \\ Y_{k,1} & \dots & Y_{k,k-r} \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} \tag{1.5}$$

with the additional condition that the matrix $(Y_{i,j})$ has rank $k-r$.

The projection of $\mathcal{V}_r$ onto the affine space with coordinates $(\mathbf{X})$ is V_r. For

$$\mathbf{u} = (u_{1,1}, \dots, u_{k-r,k}) \in \mathbb{K}^{k(k-r)},$$

we define the variety $\mathcal{V}'_{r,\mathbf{u}}$ as the intersection of $\mathcal{V}_r$ and the complex solutions of

$$\begin{bmatrix} u_{1,1} & \dots & u_{1,k} \\ \vdots & & \vdots \\ u_{k-r,1} & \dots & u_{k-r,k} \end{bmatrix} \cdot \begin{bmatrix} Y_{1,1} & \dots & Y_{1,k-r} \\ \vdots & & \vdots \\ Y_{k,1} & \dots & Y_{k,k-r} \end{bmatrix} = \mathrm{Id}_{k-r} \tag{1.6}$$

Proposition 1.100. *For any $r \in \{0, \dots, k-1\}$, the varieties V_r and $\mathcal{V}'_{r,\mathbf{u}}$ are birational.*

Proof. We need to define a morphism $f : W \longrightarrow \mathcal{V}'_{r,\mathbf{u}}$ with W a non-empty Zariski-open subset of V_r, and such that f is inverse to the projection $\mathcal{V}'_{r,\mathbf{u}} \longrightarrow V_r$ onto the affine space with coordinates $(\mathbf{X})$. Let W be the open subset of V_r defined as the non-vanishing locus of the top-left r-minor of M. Consider the block decompositions, where A, $Y_{(1)}$ and $U_{(1)}$ are $r \times r$ matrices:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad Y = \begin{bmatrix} Y_{(1)} \\ Y_{(2)} \end{bmatrix} \quad U = \begin{bmatrix} U_{(1)} & U_{(2)} \end{bmatrix}.$$

Over W, A is invertible, let $\Delta = \det(A)$. Let M/A be the Schur complement of A in M, Equations (1.5) and (1.6) can be rewritten as

$$\begin{bmatrix} \Delta \mathrm{Id}_r & A^{-1}B \\ 0 & M/A \\ U_{(1)} & U_{(2)} \end{bmatrix} \cdot \begin{bmatrix} Y_{(1)} \\ Y_{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mathrm{Id}_r \end{bmatrix}$$

We may restrict to the open subset of $\mathcal{V}'_{r,\mathbf{u}}$ where $Y_{(2)}$ is invertible, then eliminating $Y_{(1)}$ yields that

$$\begin{cases} Y_{(2)} = (U_{(2)} - U_{(1)}A^{-1}B)^{-1} \\ Y_{(1)} = \frac{-1}{\Delta} A^{-1}BY_{(2)} \end{cases}$$

which defines the wanted morphism $W \longrightarrow \mathcal{V}'_{r,\mathbf{u}}$. □

Chapter 2

Gröbner bases

This chapter deals with Gröbner bases, and gives definitions and properties which will be used throughout the thesis. It is a powerful tool for solving polynomial systems, and is defined as a particular set of generators of a polynomial ideal. This set of generators depends on the choice of an order on the monomials, and this choice has consequences on what information does the basis yield, and how complicated it is to compute the basis.

This chapter is organized as follows: in Section 2.1, we define monomial orderings, and list a few examples which shall be useful in the rest of the thesis. In Section 2.2, we recall the definition of a Gröbner basis and a normal form, and in Section 2.3 we examine two common applications of Gröbner bases (namely listing all the solutions of a system with a finite-number of solutions and computing equations of the Zariski-closure of the projection of a variety) and we use these examples to illustrate how each application leads to the choice of a specific monomial order.

Then we move on to more algorithmic considerations. In Section 2.4, we describe 3 algorithms for computing Gröbner bases: the historical algorithm from Buchberger, the more modern F_5 algorithm through its variant Matrix-F_5, and the FGLM algorithm, for change of order. Section 2.5 concludes this chapter with complexity results for algorithms Matrix-F_5 and FGLM.

This chapter, like the previous one, is purely bibliographical.

2.1. Monomial orderings

2.1.1. Definition

Gröbner bases generalize euclidean division to multivariate polynomials, and gaussian elimination to nonlinear polynomials. In both cases, we eliminate some monomials until we reach a normal form: in the univariate case, we eliminate high degree monomials; in the linear case, we choose an order on the indeterminates, and we eliminate them in that order.

Definition 2.1 (Monomial ordering)**.** Let $A = \mathbb{K}[\mathbf{X}]$ be a polynomial algebra with n indeterminates; let $\mathcal{M}$ be the set of its monomials. A *monomial ordering* is a total ordering $<_{\text{mon}}$ on $\mathcal{M}$ such that:

$$\forall x \in \mathcal{M}, 1 <_{\text{mon}} x$$
$$\forall x, y, z \in \mathcal{M}, x <_{\text{mon}} y \implies xz <_{\text{mon}} yz$$

Equivalently, a monomial ordering can be defined on the set of exponents $\mathbb{N}^n$, in which case the requirements are that:

$$\forall v \in \mathbb{N}^n, 0 <_{\text{mon}} v$$
$$\forall u, v, w \in \mathbb{N}^n, u <_{\text{mon}} v \implies u + w <_{\text{mon}} v + w$$

Given a monomial ordering, we may define:

Definition 2.2 (Leading monomial, leading term, leading coefficient)**.** Let $f \in \mathbb{K}[\mathbf{X}]$, and let $<_{\text{mon}}$ be a monomial ordering. Then

- the *leading monomial* $\mathsf{LM}_{<_{\text{mon}}}(f)$ of f is the largest monomial of its support ;
- the *leading coefficient* $\mathsf{LC}_{<_{\text{mon}}}(f)$ of f is the coefficient of $\mathsf{LM}(f)$;
- the *leading term* $\mathsf{LT}_{<_{\text{mon}}}(f)$ of f is the product $\mathsf{LC}_{<_{\text{mon}}}(f)\mathsf{LM}_{<_{\text{mon}}}(f)$.

Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, the *leading ideal* (or *initial ideal*) of I is the ideal $\langle \mathsf{LT}_{<_{\text{mon}}}(I) \rangle$ generated by the leading terms of all elements of I.

For all these notions, we shall omit the subscripted monomial order when clear by the context.

In the multivariate nonlinear case, unlike the linear and the univariate cases, the choice of an order on the monomials is not unique.

In the rest of this section, we give several examples of monomial orderings, with basic algebraic properties. We shall examine them again in Section 2.3, from the angle of applications. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n)$ be two exponents in $\mathbb{N}^n$.

2.1.2. Lexicographical ordering

The lexicographical ordering is maybe the most "natural" monomial order: it is the dictionary order for the "words" α and β.

Definition 2.3. The *lexicographical* ordering Lex is defined as:

$$\alpha <_{\text{lex}} \beta \iff \alpha_1 = \beta_1, \ldots, \alpha_{j-1} = \beta_{j-1}, \alpha_j < \beta_j$$

Its main property is the following:

Proposition 2.4. *Let $f \in \mathbb{K}[\mathbf{X}]$. If $\mathsf{LT}_{<_{lex}}(f) \in \mathbb{K}[X_i, \ldots, X_n]$ for $i < n$, then $f \in \mathbb{K}[X_i, \ldots, X_n]$.*

Proof. Let $(\alpha_1, \ldots, \alpha_n)$ be the exponents of $\mathsf{LT}(f)$ for the lexicographical order. If $\mathsf{LT}(f) \in \mathbb{K}[X_i, \ldots, X_n]$, then $\alpha_1 = \cdots = \alpha_{i-1} = 0$. Any monomial $\mathbf{X}^\beta$ such that $\beta_j > 0$ with $j < i$ is larger than $\mathbf{X}^\alpha$ for the lexicographical order. □

We shall see in Section 2.3.1 that the lexicographical ordering is particularly useful for solving zero-dimensional systems.

2.1.3. Graded reverse-lexicographical ordering

The next monomial orders that we present are degree orders: the monomials are ordered first by degree, then according to some other monomial order. We shall see that these orders are particularly suitable for working with homogeneous polynomials.

Definition 2.5. The *graded reverse-lexicographical* (or *degree reverse lexicographical*, or in short DRL) ordering GRevLex is defined as:

$$\alpha <_{\text{grevlex}} \beta \iff \begin{cases} \deg(\mathbf{X}^\alpha) < \deg(\mathbf{X}^\beta) \\ \text{or} \\ \deg(\mathbf{X}^\alpha) = \deg(\mathbf{X}^\beta) \text{ and } \alpha_i > \beta_i\,,\ \alpha_{i+1} = \beta_{i+1}\,, \ldots,\ \alpha_n = \beta_n \end{cases}$$

This definition extends straightforwardly to a weighted setting:

Definition 2.6. Let $W \in \mathbb{N}^n$ be a system of weights.The *weighted graded reverse-lexicographical* ordering W-GRevLex is defined as:

$$\alpha <_{W\text{-grevlex}} \beta \iff \begin{cases} \deg_W(\mathbf{X}^\alpha) < \deg_W(\mathbf{X}^\beta) \\ \text{or} \\ \deg_W(\mathbf{X}^\alpha) = \deg_W(\mathbf{X}^\beta) \text{ and } \alpha_i > \beta_i\,,\ \alpha_{i+1} = \beta_{i+1}\,, \ldots,\ \alpha_n = \beta_n \end{cases}$$

When working with homogeneous polynomials, the leading term for these orders contains a lot of information on the remaining terms. The next elementary propositions shall be useful in Section 2.5.3, in order to refine the complexity analyses for Algorithm F_5.

Proposition 2.7. *Let $f \in \mathbb{K}[\mathbf{X}]$ be a homogeneous* (resp. *W-homogeneous) polynomial. If $X_i^k \mid \mathsf{LT}_{<_{grevlex}}(f)$* (resp. *$X_i^k \mid \mathsf{LT}_{<_{W\text{-}grevlex}}(f)$), then $f \in \langle X_i^k, X_{i+1}, \ldots, X_n\rangle$. In particular, if $\mathsf{LT}(f)$ is divisible by X_n, then f is divisible by X_n.*

Proof. It suffices to prove it in the weighted case, the total degree case is obtained by setting all weights to 1.

First assume that $\mathsf{LT}(f) = X_1^{\alpha_1} \cdots X_i^{\alpha_i}$, or in other words that i is the largest index of a variable dividing $\mathsf{LT}(f)$. Let μ be a monomial with the same W-degree, and such that $\mu \notin \langle X_i^{\alpha_i}, X_{i+1}, \ldots, X_n\rangle$, this means that

$$\mu = X_1^{\beta_1} \cdots X_i^{\beta_i} \text{ with } \beta_i < \alpha_i,$$

so $\mu <_{\text{lex}} \mathsf{LT}(f)$, and thus $\mu >_{W\text{-grevlex}} \mathsf{LT}(f)$, and it may not appear in the support of f.

If i is not the largest index of a variable dividing $\mathsf{LT}(f)$, then $\mathsf{LT}(f)$ is divisible by X_j for some $j > i$, and from the above

$$f \in \langle X_j, \ldots, X_n\rangle \subset \langle X_i^{\alpha_i}, X_{i+1}, \ldots, X_j, \ldots, X_n\rangle.$$

□

For $m \in \{1, \dots, n\}$, let θ_m be the morphism evaluating $X_{m+1}, \dots, X_n$ to 0:

$$\begin{array}{rccll} \theta_m & : & \mathbb{K}[\mathbf{X}] & \longrightarrow & \mathbb{K}[\mathbf{X}] \\ & & X_i & \longmapsto & X_i \quad (i \leq m) \\ & & X_i & \longmapsto & 0 \quad (i > m) \end{array}$$

Corollary 2.8. *Let $f \in \mathbb{K}[\mathbf{X}]$ be a homogeneous (*resp. *W-homogeneous) polynomial, and let $\prec$ be the* GRevLex *(*resp. *W-*GRevLex*) order. If $m \in \{1, \dots, n\}$ is such that $\theta_m(f) \neq 0$, then*

$$\mathsf{LT}_\prec(f) = \mathsf{LT}_\prec(\theta_m(f))$$

Proof. The evaluation $\theta_m(f)$ is nonzero if and only if $f \notin \langle X_{m+1}, \dots, X_n \rangle$. So if $\theta_m(f) \neq 0$, $\mathsf{LT}(f)$ is not divisible by $X_{m+1}, \dots, X_n$, and

$$\mathsf{LT}(\theta_m(f)) = \theta_m(\mathsf{LT}(f)) = \mathsf{LT}(f).$$

□

Proposition 2.9 ([BFS14, Prop. 7])**.** *Let $(f_1, \dots, f_m)$ be a homogeneous (resp. W-homogeneous) polynomial system in Noether position with respect to the variables $X_{m+1}, \dots, X_n$, let $\prec$ the* GRevLex *(*resp. *W-*GRevLex*) order, and $I = \langle f_1, \dots, f_m \rangle$. Then:*

$$\mathsf{LT}_\prec(I) = \mathsf{LT}_\prec(\theta_m(I)) \cdot \langle X_{m+1}, \dots, X_n \rangle.$$

Proof. It suffices to prove it in the weighted case, the total degree case is obtained by setting all weights to 1.

From Corollary 2.8, for the W−GRevLex ordering, $\mathsf{LT}(\theta_m(I)) \subset \mathsf{LT}(I)$, this implies the reverse inclusion

$$\mathsf{LT}(I) \supset \mathsf{LT}(\theta_m(I)) \cdot \langle X_{m+1}, \dots, X_n \rangle.$$

Conversely, let $f \in I$ be a W-homogeneous polynomial, write $\mathsf{LT}(f) = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. We want to prove that there exists $g \in I$ such that $\mathsf{LT}(g) = X_1^{\alpha_1} \cdots X_m^{\alpha_m}$. Let i be the largest index of a variable dividing $\mathsf{LT}(f)$ (that is the largest integer such that $\alpha_i \neq 0$). If $i \leq m$, there is nothing to prove.

Otherwise, from Proposition 2.7, this implies that $f \in \langle X_i^{\alpha_i}, X_{i+1}, \dots, X_i \rangle$, which we may write

$$f = X_i^{\alpha_i} g_i + \sum_{j=i+1}^{n} X_j g_j$$

with g_j ($j \leq i$) W-homogeneous and $0 \neq g_i \in \mathbb{K}[X_1, \dots, X_i]$ (all monomials divisible by X_j, for $j > i$, are pushed in the term $X_j g_j$). This implies that

$$X_i^{\alpha_i} g_i \in I + \langle X_{i+1}, \dots, X_n \rangle. \tag{2.1}$$

By Proposition 1.44, since $(f_1, \dots, f_m)$ is in Noether position, $(f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ is a regular sequence, and by Corollary 1.39, $(f_1, \dots, f_m, X_n, \dots, X_{m+1})$ is also regular. Equation (2.1) means that $X_i^{\alpha_i}$ is a zero divisor in $\mathbb{K}[\mathbf{X}]/\langle f_1, \dots, f_m, X_n, \dots, X_{i+1} \rangle$, so by regularity,

$$g_i \in \langle f_1, \dots, f_m, X_n, \dots, X_{i+1} \rangle = I + \langle X_{i+1}, \dots, X_n \rangle.$$

Since $g_i \in \mathbb{K}[X_1, \dots, X_i]$, we deduce that $g_i \in I$, and its leading term is $X_1^{\alpha_1} \cdots X_{i-1}^{\alpha_{i-1}}$. Repeating the process until $i \leq m$, we obtain the wanted polynomial. □

2.1.4. Elimination orders

The last order that we present is an elimination ordering.

Definition 2.10. Let k be a integer such that $1 \leq k \leq n$, the *k'th elimination ordering* $\textsc{Elim}_k$ is defined as:

$$\alpha <_{\textsc{Elim}_k} \beta \iff \begin{cases} (\alpha_1, \ldots, \alpha_k) <_{\text{grevlex}} (\beta_1, \ldots, \beta_k) \\ \text{or} \\ (\alpha_1, \ldots, \alpha_k) = (\beta_1, \ldots, \beta_k) \text{ and } (\alpha_{k+1}, \ldots, \alpha_n) <_{\text{grevlex}} (\beta_{k+1}, \ldots, \beta_n) \end{cases}$$

The name comes from the fact that this ordering can be used to eliminate variables.

Definition 2.11. Let $<_{\text{mon}}$ be a monomial ordering, we say that $<_{\text{mon}}$ *eliminates* the first k variables if and only if for any $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $(\beta_{k+1}, \ldots, \beta_n) \in \mathbb{N}^{n-k}$,

$$(0, \ldots, 0, \beta_{k+1}, \ldots, \beta_n) <_{\text{mon}} (\alpha_1, \ldots, \alpha_n)$$

Equivalently, for $f \in \mathbb{K}[\mathbf{X}]$, if $\mathsf{LT}_{<_{\text{mon}}}(f) \in \mathbb{K}[X_{k+1}, \ldots, X_n]$, then $f \in \mathbb{K}[X_{k+1}, \ldots, X_n]$.

From Proposition 2.4, the lexicographical order eliminates the first k variables for all k. On the other hand, the graded reverse lexicographical ordering does not eliminate any group of variables in general (there may be monomials with lower degree in the trailing terms). The elimination orderings are in some sense an intermediate ground between the lexicographical and the graded reverse-lexicographical orderings:

Proposition 2.12. *Let $k \in \{1, \ldots, n\}$, the k'th elimination ordering eliminates the first k variables.*

Proof. Immediate consequence of the definition. □

2.2. Gröbner bases: definition

We can now define Gröbner bases.

Definition 2.13 (Gröbner basis)**.** Let I be an ideal of $\mathbb{K}[\mathbf{X}]$. A *Gröbner basis* of I is a subset $G = \{g_1, \ldots, g_k\} \subset I$ such that

$$\langle \mathsf{LT}(g_1), \ldots, \mathsf{LT}(g_k) \rangle = \langle \mathsf{LT}(I) \rangle.$$

This definition can be rephrased as follows: a Gröbner basis of I is a finite subset $G \subset I$ and such that

$$\forall\, f \in I, f \neq 0 \implies \exists g \in G, \mathsf{LT}(g) \mid \mathsf{LT}(f).$$

The fundamental property of Gröbner bases is that they define a normal form:

Proposition/Definition 2.14 (Normal Form application, [BW93, Th. 5.35])**.** Let I be an ideal of $\mathbb{K}[\mathbf{X}]$ and let G be a Gröbner basis of I. Let $f \in \mathbb{K}[\mathbf{X}]$ be a polynomial. There exists a *unique* polynomial $\mathsf{NF}(f) \in \mathbb{K}[\mathbf{X}]$ such that:

- $f - \mathsf{NF}(f) \in \langle G \rangle$ (we say that f *reduces* to $\mathsf{NF}(f)$ modulo G, written $f \xrightarrow{G} \mathsf{NF}(f)$);
- For all μ in the support of $\mathsf{NF}(f)$, μ is not divisible by any $\mathsf{LT}(g)$, $g \in G$.

The polynomial $\mathsf{NF}(f)$ is called the *normal form* of f modulo G. If $f = \mathsf{NF}(f)$, we say that f is *in normal form*. Otherwise, f is *reducible* modulo G.

If $\mathsf{LT}(f)$ is reducible modulo G, we say that f is *head-reducible* modulo G. If f is reducible modulo G but $\mathsf{LT}(f)$ is irreducible, we say that f is *tail-reducible*.

This notion allows to prove that a Gröbner basis of I is a set of generators of I.

Proposition 2.15. *Let G be a Gröbner basis of I, then $\langle G \rangle = I$.*

Proof. By definition, $\langle G \rangle \subset I$. Let $f \in I$, $f - \mathsf{NF}(f) \in \langle G \rangle \subset I$, so $\mathsf{NF}(f) \in I$. Assume that $\mathsf{NF}(f) \neq 0$, then by definition of a Gröbner basis, $\mathsf{LT}(\mathsf{NF}(f))$ is divisible by $\mathsf{LT}(g)$ for some $g \in G$, thus contradicting the definition of $\mathsf{NF}(f)$. Hence $\mathsf{NF}(f) = 0$ and $f \in \langle G \rangle$. □

The normal form application has a lot more properties which we list now.

Proposition 2.16 ([**CLO07**])**.** *Let $f, g \in \mathbb{K}[\mathbf{X}]$ and let NF be a normal form application for a Gröbner basis G of some ideal I. Then:*

1. *$\mathsf{NF}(f)$ is in normal form with respect to G;*
2. *$\mathsf{NF}(f) = 0$ if and only if $f \in I$;*
3. *$\mathsf{NF}(f) = \mathsf{NF}(g)$ if and only if $f - g \in I$;*
4. *$\mathsf{NF}(f + g) = \mathsf{NF}(f) + \mathsf{NF}(g)$;*
5. *$\mathsf{NF}(fg) = \mathsf{NF}(\mathsf{NF}(f)\mathsf{NF}(g))$;*

In particular, NF defines a surjective ring morphism $\mathbb{K}[\mathbf{X}] \longrightarrow \mathbb{K}[\mathbf{X}]/I$.

Proof. Item 1 is a consequence of the unicity of the normal form. Item 2 is a consequence of the proof of Proposition 2.15.

For item 3, write

$$f - g = (f - \mathsf{NF}(f)) + (\mathsf{NF}(f) - \mathsf{NF}(g)) + (\mathsf{NF}(g) - g),$$

so $f - g \in I$ if and only if $\mathsf{NF}(f) - \mathsf{NF}(g) \in I$. In particular, if $\mathsf{NF}(f) - \mathsf{NF}(g) = 0$, then $f - g = 0$.

Conversely, the support of $\mathsf{NF}(f) - \mathsf{NF}(g)$ is a subset of the union of the supports of $\mathsf{NF}(f)$ and $\mathsf{NF}(g)$, so all monomials of $\mathsf{NF}(f) - \mathsf{NF}(g)$ are irreducible modulo G. So $\mathsf{NF}(f) - \mathsf{NF}(g)$ is in normal form, and so if $\mathsf{NF}(f) - \mathsf{NF}(g) \in I$, then it is zero.

Items 4 and 5 are proved similarly: since

$$\mathsf{NF}(f + g) - \mathsf{NF}(f) - \mathsf{NF}(g) = (\mathsf{NF}(f + g) - f - g) + (f - \mathsf{NF}(f)) + (g - \mathsf{NF}(g)) \in I,$$

the same argument involving the support shows that

$$\mathsf{NF}(f + g) - \mathsf{NF}(f) - \mathsf{NF}(g) = 0.$$

And since

$$fg - \mathsf{NF}(f)\mathsf{NF}(g) = f(g - \mathsf{NF}(g)) - (f - \mathsf{NF}(f))\mathsf{NF}(g) \in I,$$

taking the normal form of both terms of the left hand side yields that

$$\mathsf{NF}(fg) - \mathsf{NF}(\mathsf{NF}(f)\mathsf{NF}(g)) = 0. \qquad \square$$

Item 2 is the main property of Gröbner bases: the normal form application that they define give a practical way of testing whether any polynomial belong to the ideal.

The last two propositions of this section are existence statements for Gröbner bases. We shall describe algorithms computing Gröbner bases later, which will give a constructive proof of these results.

Proposition 2.17 ([CLO07, Ch. 2, sec. 5, cor. 6])**.** *Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, then I admits a Gröbner basis for any given monomial order.*

Gröbner bases are not unique in general, since we may add polynomials from the ideal to any Gröbner basis and still obtain a Gröbner basis. However, with additional properties, unicity can be ensured.

Definition 2.18. Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, and let G be a Gröbner basis of I. We say that G is *reduced* if the following two conditions are satisfied:

1. all leading coefficients of polynomials of G are 1;
2. all polynomials in G are in normal form with respect to the others.

Proposition 2.19 ([CLO07, Ch. 2, sec. 7, prop. 6])**.** *Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, then I admits a unique reduced Gröbner basis for any given monomial order.*

2.3. Applications

What information one can obtain from a Gröbner basis depends on the monomial order that is used. The most useful orders in that regard are the lexicographical order and the elimination orders.

2.3.1. Zero-dimensional systems

Recall that zero-dimensional systems are systems with a finite number of solutions. A first property is that we can check this property using any Gröbner basis.

Proposition 2.20. *Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, and let G be a Gröbner basis of I. The ideal I is zero-dimensional if and only if for all $X \in \{\mathbf{X}\}$, there exists a polynomial in G with leading term X^α for some α.*

Proof. Recall that I is zero-dimensional if and only if $\mathbb{K}[\mathbf{X}]/I$ is a $\mathbb{K}$-vector space with finite dimension. Let

$$M_I = \{\mu \text{ monomial of } \mathbb{K}[\mathbf{X}] \mid \forall\, g \in G, \mathsf{LT}(g) \nmid \mu\},$$

elements of M_I are all linearly independent modulo I. Hence M_I is finite.

Let $X \in \{\mathbf{X}\}$, there exists $\beta \in \mathbb{N}$ such that $X^\beta \notin M_I$, and so there exists $\alpha \in \mathbb{N}$, $\alpha \leq \beta$ such that X^α is the leading term of an element of G. □

The lexicographical order allows one to solve generic zero-dimensional systems.

Definition 2.21 (Shape position)**.** Let I be an ideal in $\mathbb{K}[\mathbf{X}]$, and let G be the reduced lexicographical Gröbner basis of I. We say that I is in *shape position* if G has the following shape:

$$\begin{cases} g_1(X_1, \ldots, X_n) & = X_1 + r_1(X_n) \\ g_2(X_2, \ldots, X_n) & = X_2 + r_2(X_n) \\ & \vdots \\ g_{n-1}(X_{n-1}, X_n) & = X_{n-1} + r_{n-1}(X_n) \\ g_n(X_n) & = X_n^{\deg(I)} + r_n(X_n) \end{cases}$$

In particular, if I is in shape position, it is zero-dimensional, and its $\deg(I)$ roots can be recovered by solving the univariate polynomial g_n.

This property is generic in the following sense:

Lemma 2.22 (Shape lemma, [GM89])**.** *Up to a generic change of coordinates, any zero-dimensional ideal is in shape position.*

This makes lexicographical Gröbner bases a very powerful tool for solving zero-dimensional systems.

2.3.2. Positive-dimensional systems and eliminations

Gröbner bases can also be used to work with positive-dimensional systems. In particular, elimination bases allow to "compute" projections: the projection of an algebraic variety onto an affine subspace is not in general an algebraic variety, but its Zariski closure is, and one can compute equations defining this algebraic set by polynomial elimination.

Proposition 2.23 ([CLO07, Chap. 3, Sec. 1, Th. 2])**.** *Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, let $k \in \{1, \ldots, n\}$, and let G be a Gröbner basis of I for an order eliminating the first k variables (for example $\textsc{Elim}_k$ or $\textsc{Lex}$).*

Then the elimination basis $G_k := G \cap \mathbb{K}[X_{k+1}, \ldots, X_n]$ *is a Gröbner basis of $I \cap \mathbb{K}[X_{k+1}, \ldots, X_n]$. Geometrically, G_k is a system defining the Zariski closure of the projection of $V(I)$ onto the coordinates $(X_{k+1}, \ldots, X_n)$.*

Remark 2.24. In [CLO07, Chap. 3, Sec. 1, Th. 2], the theorem is written for the lexicographical order, but the proof only uses the fact that this order eliminating the first k variable.

Remark 2.25. The shape position is an example of successive eliminations: with the notations of Definition 2.21, for any $i \in 1, \dots, n$, the polynomials $\{g_i, \dots, g_n\}$ generate the ideal $I \cap \mathbb{K}[X_i, \dots, X_n]$. In particular, the last polynomial g_n generates $I \cap \mathbb{K}[X_n]$.

Geometrically, a zero-dimensional ideal I has finitely-many solutions. Their projection on the X_n line is a finite set, and I being in shape position means that each of these roots x_n is the projection of only 1 point $(x_1, \dots, x_n) \in V(I)$ onto the X_n line, and this point is given by

$$(-r_1(x_n), \dots, -r_n(x_n), x_n).$$

Another example of how elimination techniques can be useful is saturation:

Proposition 2.26 ([CLO07, Sec. 4.4, Ex. 9])**.** *Let I be an ideal of $\mathbb{K}[\mathbf{X}]$, let $F = (f_1, \dots, f_m)$ be a system of generators of I, and let $g \in \mathbb{K}[\mathbf{X}]$. Consider the* saturated system

$$F_{\text{sat}} \coloneqq (f_1, \dots, f_m, U \cdot g - 1) \subset \mathbb{K}[U, X_1, \dots, X_n].$$

Let G_{sat} be a Gröbner basis of $\langle F_{\text{sat}} \rangle$ eliminating the first variable U. Then G_{sat} is a Gröbner basis of the saturated ideal $(I : g^\infty)$ *defined as*

$$(I : g^\infty) \coloneqq \{f \in \mathbb{K}[\mathbf{X}] \mid \exists\, n \in \mathbb{N}, g^n f \in I\}.$$

Geometrically, G_{sat} is a system defining the Zariski closure of the complementary of $V(\langle g \rangle)$ in $V(I)$.

2.4. Algorithms

Many algorithms have been developed to compute Gröbner bases, we will present three examples here.

2.4.1. A pairs algorithm: Buchberger's algorithm

The first algorithm for computing Gröbner bases was Buchberger's algorithm [Buc76]. Its building block is pairwise reductions.

Definition 2.27 (S-polynomial)**.** Let $f, g \in \mathbb{K}[\mathbf{X}]$, the *$S$-polynomial* of f and g is defined as

$$S\text{-pol}(f, g) \coloneqq \operatorname{lcm}(\mathsf{LM}(f), \mathsf{LM}(g)) \left(\frac{f}{\mathsf{LM}(g)} - \frac{g}{\mathsf{LM}(f)} \right)$$

In particular, if f is not in normal form with respect to g, $S\text{-pol}(f, g)$ is a reduction of f modulo g:

$$S\text{-pol}(f, g) = f - \frac{\operatorname{lcm}(\mathsf{LM}(f), \mathsf{LM}(g))}{\mathsf{LM}(f)} \cdot g.$$

This allows for computing polynomial reductions algorithmically.

Algorithm 2.1 Reduction

Input: $f \in \mathbb{K}[\mathbf{X}], G \subset \mathbb{K}[\mathbf{X}]$

Output: $h \in \mathbb{K}[\mathbf{X}]$ such that $f \xrightarrow{G} h$ and h is in normal form with respect to G

```
h ← f
res ← 0
while h ≠ 0 do
    while ∃ g ∈ G such that LT(g) | LT(h) do
        h ← S-pol(h, g)
    end while
    res ← res + LT(h)
    h ← h − LT(h)
end while
return res
```

Remark 2.28. If G is a Gröbner basis and f is a polynomial, Reduction(f,G) is the normal form of f with respect to G.

Remark 2.29. If G is a Gröbner basis of some ideal I, let

$$G' := \{\text{Reduction}(g, G \setminus g) \mid g \in G\} \setminus \{0\}.$$

Then G' is a *reduced* Gröbner basis of I for the same monomial order.

Buchberger's algorithm uses S-polynomials and Algorithm Reduction to compute a Gröbner basis of an ideal: given a set of generators of the ideal, it creates all S-polynomials of pairs of these generators, then reduces them modulo the initial set of generators. The result is a new, larger set of generators, and it iterates these steps until all S-polynomials reduce to zero.

Algorithm 2.2 Buchberger ([Buc76])

Input: $F \subset \mathbb{K}[\mathbf{X}]$

Output: $G \subset \mathbb{K}[\mathbf{X}]$ such that G is a Gröbner basis of $\langle F \rangle$

```
G ← F
Q ← F × F
while Q ≠ ∅ do
    Pick (f, g) in Q, remove (f, g) from Q
    h ← S-pol(f, g)
    h ← Reduction(h, G)
    if h ≠ 0 then
        Q ← Q ∪ G × {h}
        G ← G ∪ {h}
    end if
end while
return G
```

Theorem 2.30 (Correctness of Buchberger's algorithm, [Buc76],[CLO07, Ch. 2, sec. 7, th. 2])**.** *Buchberger's algorithm terminates and returns a Gröbner basis of I.*

Remark 2.31. Buchberger's algorithm being correct gives another proof that Gröbner bases exist for all ideals.

This version of Buchberger's algorithm is only of theoretical interest. In practice, many pairs will eventually reduce to the zero polynomial, and thus are effectively useless for the rest of the computations. To mitigate this effect, two algorithmic criteria were given in [Buc79], filtering out some pairs leading to reductions to zero. The algorithm can also be improved by defining strategies for the choice of the next pair (for example the *sugar strategy*, see [Gio+91]) Buchberger's algorithm is currently implemented, for example in Macaulay2 [Macaulay2], Singular [Singular] or for computations over arbitrary rings in Magma[Magma] or Maple[Maple].

Buchberger's algorithm belongs to a wide family of Gröbner basis algorithms, called *pairs algorithms*: they construct a set of pairs, reduce them, and build new pairs with the result.

Other examples include the HDriven algorithm [Tra96] (Buchberger with an additional criterion based on information on the Hilbert series of the ideal, currently implemented in most systems implementing Buchberger), the F_4 algorithm [Fau99] (selecting several pairs at each step and using linear algebra instead of S-polynomials to compute the reductions, currently implemented in the library FGb [Fau10], in Magma and in Maple), and the F_5 algorithm [Fau02] (F_4 with the additional F_5 criterion, see Section 2.4.2).

2.4.2. A matrix algorithm: Matrix-F5

The complexity of the previous algorithms is difficult to estimate, because it mainly depends on the length of the queue at any given step. However, the complexity of these algorithms can be bounded by studying matrix variants, replacing S-polynomials and critical pairs with the construction and reduction of Macaulay matrices.

For example, the matrix variant of Buchberger's algorithm is Lazard's algorithm [Laz83]: at each degree d, it builds the Macaulay matrix of degree d of the system, reduces it to echelon form, and extract a new set of generators from the new matrix.

In this section, we describe a matrix variant of F_5 algorithm, called Matrix-F_5 ([BFS14]).

This algorithm works by keeping track, for each computed polynomial, of "where it came from". It uses this information to exclude some rows from the constructed matrix: the matrix built is a submatrix of the Macaulay matrix, whose rows span the same vector space as the rows of the Macaulay matrix.

Definition 2.32 (Signature)**.** Let $F = (f_1, \dots, f_m)$ be a family of homogeneous polynomials in $\mathbb{K}[\mathbf{X}]$, and let $I = \langle F \rangle$. Consider a polynomial $g \in I$ such that

$$g = g_i f_i + \sum_{j=1}^{i-1} g_j f_j$$

with $g_i \neq 0$ and i minimal for this property. The *signature* of g is the pair $(i, \mathsf{LT}(g_i))$.

Signatures are ordered using the lexicographical ordering on the two components:

$$(i, \mu) < (j, \nu) \iff \begin{cases} i < j \\ \text{or} \\ i = j \text{ and } \mu < \nu \end{cases}$$

Remark 2.33. Decomposing the g_j's into sums of monomials, the definition means that a polynomial with signature (i, μ) can be written as a linear combination of μf_i and polynomials with smaller signatures.

Lemma 2.34 (F_5 criterion, [Fau02]). *Let $F = (f_1, \ldots, f_m)$ be a family of homogeneous polynomials in $\mathbb{K}[\mathbf{X}]$, and let $I = \langle F \rangle$. Let $i \in \{1, \ldots, m\}$ and μ be a monomial of $\mathbb{K}[\mathbf{X}]$. If μ is the leading term of a polynomial in $\langle f_1, \ldots, f_{i-1} \rangle$, then the polynomial μf_i is a linear combination of polynomials with signature less than (i, μ). In other words, no polynomial in I has signature (i, μ).*

Proof. Let $\mu = \mathsf{LT}(g)$, and let $(j, \bullet)$ be the signature of g. By hypothesis, $j < i$. Write

$$g = g_j f_j + \sum_{k=1}^{j-1} g_k f_k$$

and let $r = g - \mu$. All monomials in r are smaller than μ. Write

$$\begin{aligned} \mu f_i &= g f_i - r_2 f_i \\ &= -r f_i + \sum_{k=1}^{j} g_k f_i f_k \end{aligned}$$

Hence μf_i has signature at most $(i, \mathsf{LT}(r))$ with $\mathsf{LT}(r) < \mu$. □

Unlike pairs-based algorithms which stop when the pair queue is empty, there is no obvious stopping criterion for matrix-based algorithms. This means that the algorithm takes an additional parameter along with the system, a degree $d_{\max}$ at which to stop. Formally, what the algorithm computes is a *Gröbner basis truncated to $d_{\max}$* (or a *$d_{\max}$-Gröbner basis*), that is a set of polynomials containing the polynomials of the reduced Gröbner basis of I of degree at most $d_{\max}$.

In Algorithm 2.3, EchelonForm denotes a routine performing Gaussian elimination on a matrix, with restrictions:

- rows are not exchanged;
- rows are only reduced by rows above them.

These conditions will be used to ensure that any polynomial is only reduced by polynomials with smaller signatures.

The notation $\mathrm{row}(M, i)$ is a shorthand for the polynomial associated with the i'th row of the matrix M. When talking about matrices, the notation $A \cup B$ means that, both matrices having the same number of columns, we append matrix B to the bottom of matrix A.

The algorithm proceeds by looping over the degree d, and then over the polynomials f_i, and then over the monomials μ with degree $d - d_i$, as if to build the Macaulay matrix with rows μf_i. For each row μf_i, it tests if it satisfies the F_5-criterion (line 11), and if not, it computes its signature and adds the row to the matrix.

Theorem 2.35 (Correctness of algorithm Matrix-F_5, [Fau02; BFS14]). *Algorithm Matrix-F_5 terminates and computes a $d_{\max}$-Gröbner basis of F. Furthermore, if F is a regular sequence, then no rows are reduced to zero in the echelon forms.*

Algorithm 2.3 Matrix-F_5 ([BFS14])

Output: $F \subset \mathbb{K}[\mathbf{X}]$ *homogeneous* system, $d_{\max} \in \mathbb{N}$ an implicit monomial ordering
Input: $G \subset \mathbb{K}[\mathbf{X}]$, truncated Gröbner basis of $\langle F \rangle$ at degree $d_{\max}$

$G \leftarrow \varnothing$
for $d = 0$ **to** $d_{\max}$ **do**
  $N_d \leftarrow$ number of monomials of degree d
  $M_{d,0}, \tilde{M}_{d,0} \leftarrow$ matrix with 0 rows and N_d columns
  for $i = 1$ **to** m **do**
    $M_{d,i} \leftarrow M_{d,i-1}$
    if $d = d_i$ **then**
      $M_{d,i} \leftarrow M_{d,i} \cup$ row f_i with signature $(i, 1)$
    else if $d > d_i$ **then**
      for all μ monomial of degree $d - d_i$ **do**
        if μ is not leading term of a row in $\tilde{M}_{d-d_i,i-1}$ **then**
          $j \leftarrow$ largest j such that X_j divides μ
          $\mu' \leftarrow \mu / X_j$
          $\tilde{f} \leftarrow$ row of $\tilde{M}_{d-1,i}$ with signature (i, μ')
          $M_{d,i} \leftarrow M_{d,i} \cup$ row $X_j \tilde{f}$ with signature (i, μ)
        end if
      end for
    end if
  end for
  $\tilde{M}_{d,i} \leftarrow \mathsf{EchelonForm}(M_{d,i})$
  $R \leftarrow$ number of rows in $M_{d,i}$
  $G \leftarrow G \cup \{\mathsf{row}(\tilde{M}_{d,i}, j) \mid j \in \{1, \ldots, R\}, \mathsf{row}(\tilde{M}_{d,i}, j) \neq \mathsf{row}(M_{d,i}, j)\}$
end for
return G

Proof. For all $i \in \{1, \dots, m\}$, let $F_i = \{f_1, \dots, f_i\}$ and $I_i = \langle F_i \rangle$.

We will prove by recurrence over (d, i) that the rows of $M_{d,i}$ span $(I_i)_d$. Note that for any (i, d) such that it is true, after Gaussian elimination, the rows of $\tilde{M}_{d,i}$ still span $(I_i)_d$, and the leading terms of these rows are the leading terms of $(I_i)_d$.

The base cases $d = \min(d_j)$ and $i = 0$ for any d are clear.

Assume that $d > \min(d_j)$ and $i > 0$. The matrix $M_{d,i}$ is a submatrix of the Macaulay matrix Mac_d of F_i. More precisely, it is obtained by removing from Mac_d those rows failing to match the F_5 criterion.

By induction hypothesis, the rows with signature $(j, \bullet)$, $j < i$ span $(I_{i-1})_d$, so we need only check those rows with signature $(i, \bullet)$. If $M_{d,i}$ does not span $(I_i)_d$, then there exists a polynomial

$$f = \sum_{j=1}^{i} g_j f_j$$

which does not lie in $\mathrm{Vect}(M_{d,i})$. Let $\mu = \mathsf{LT}(g_j)$, and assume that f is such that μ is minimal. Write $g_i = \mu + r_i$, all monomials in the support of r_i are smaller than μ, and by minimality of μ

$$r \coloneqq r_i f_i + \sum_{j=1}^{i-1} g_j f_j \in \mathrm{Vect}(M_{d,i})$$

Furthermore, since signatures are inserted in increasing order, r lies in the vector space spanned by all rows inserted before considering signature (i, μ).

If the row with signature (i, μ) is inserted in the matrix, then f is the sum of this line and r, and so it lies in $\mathrm{Vect}(M_{d,i})$. So the signature (i, μ) was rejected by the F_5 criterion, which implies that the row μf_i is a linear combination of rows with signature less than (i, μ), and so again it lies in $\mathrm{Vect}(M_{d,i})$. So we have reached a contradiction, and we conclude that

$$\mathrm{Vect}(M_{d,i}) = (I_i)_d.$$

For the second part of the statement, assume that there is a row reduced to zero in a run of $\mathsf{Matrix\text{-}F}_5$, and let (i, μ) be the signature of such a row. It means that there exist polynomials $g_1, \dots, g_i$ such that

$$\sum_{j=1}^{i} g_j f_j = 0,$$

and so

$$g_i f_i \in \langle f_1, \dots, f_{i-1} \rangle.$$

If F is regular, this implies that $g_i \in I_{i-1}$, and so $\mu = \mathsf{LT}(g_i) \in \mathsf{LT}(I_{i-1})$. Thus this row is rejected by the F_5 criterion, and we reached a contradiction. □

2.4.3. Matrix algorithms in the affine case

In the previous section, we described a matrix algorithm computing Gröbner bases for homogeneous systems. The idea was to build a full-rank submatrix of the Macaulay matrix of the system, and reduce it to echelon form.

Macaulay matrices can be defined in an affine setting, in the following way:

Definition 2.36. If $F = (f_1, \dots, f_m) \in \mathbb{K}[\mathbf{X}]^m$ is any polynomial system (not necessarily homogeneous), the Macaulay matrix of F is defined as the Macaulay matrix of the homogenized system F^h. Equivalently, it is the matrix of a basis of the set of products μf_i of degree *at most* d.

In general, this set of products is *not* a basis of the set of polynomials in $\langle F \rangle$ with degree at most d.

Example 2.37. In $\mathbb{K}[X, Y]$, let $f_1 = XY + Y$ and $f_2 = X$. At total degree 1, the only polynomial in the Macaulay matrix is $f_2 = X$. However, $f_1 - Yf_2 = Y$ is also a degree 1 polynomial in the ideal.

Matrix-based Gröbner basis algorithms can be adapted to the affine case in a way similar to the construction of the affine Macaulay matrix: compute F^h, run the algorithm on this homogeneous system, and dehomogenize the result. But for the same reason, Matrix-$\mathsf{F}_5(F^h, d_{\max})$ does *not* necessarily return a $d_{\max}$-Gröbner basis of $\langle F \rangle^h$.

Example 2.38. With the system of Example 2.37, the output of Matrix-$\mathsf{F}_5(F, 1)$ is X, which does not generate $\langle F \rangle = \langle F \rangle^h = \langle X, Y \rangle$.

This behavior is called a degree fall:

Definition 2.39. Let F be a polynomial system and let F^h be the homogenized system (with an homogenization variable H). We say that Matrix-F_5 has a *degree fall* if it computes a row corresponding to a polynomial divisible by H^k for some k.

Ignoring degree falls is not efficient, because it requires the algorithm to reach an unnecessarily high degree. Instead, degree falls can be handled in the following way:

1. Detect them in all lines computed at a given step;
2. Reinject them into the matrix at the relevant degree

Note that in order to detect whether some polynomial f represents a degree fall, for the GRevLex order, it suffices to check whether $\mathsf{LT}(f)$ is divisible by H.

This technique mitigates the additional complexity induces by degree falls, and it is quite efficient in some cases. However, estimating the complexity of the algorithms now requires to be able to count the number of degree falls and their magnitude.

For this reason, the complexity of matrix Gröbner basis algorithms for affine systems is often evaluated under some generic assumption ensuring that there is no degree fall.

Definition 2.40. Let F be a polynomial system. We say that F is *regular in the affine sense* if the sequence $(h_1, \dots, h_m)$, where for all $i \in \{1, \dots, m\}$, h_i is the highest degree component of f_i, is regular.

Proposition 2.41. *Let $d_1, \dots, d_m \in \mathbb{N}_{>0}$. Sequences regular in the affine sense are generic amongst sequence of polynomials of respective degree less than d_i.*

Proof. There exists a non-empty Zariski-open subset U of the space $A_=$ of families of *homogeneous* polynomials of respective degree d_i such that U is contained in the set of regular sequences. The space of all families of arbitrary polynomials of respective degree d_i is the product of $A_=$ and the space $A_<$ of all polynomials of respective degree less than d_i, and the non-empty Zariski-open set $U \times A_<$ is contained in the set of sequences regular in the affine sense. □

Proposition 2.42 ([Bar04])**.** *Let F be a polynomial system, assume that F is regular in the affine sense. Let $d_{\max} \in \mathbb{N}$, let $G_h = \mathit{Matrix}\text{-}\mathsf{F}_5(F^h, d_{\max})$ and let $G = G_h^a$. Then Matrix-F_5 computes G_h without any degree fall, and G is a $d_{\max}$-Gröbner basis of $\langle F \rangle$;*

Furthermore, assume that we are using the GRevLex monomial order, and let d_{reg} be the degree of regularity of the sequence F_h of the highest degree components of F. Then if $d_{\max} \geq d_{\mathrm{reg}}$, then G is a Gröbner basis of $\langle F \rangle$.

Proof. A degree fall is a reduction to zero of the highest degree components. Since by hypothesis these highest degree components form a regular sequence, the F_5 Criterion ensures that no such reduction happen.

If there is no degree fall, the rows of the F_5 matrix at any degree d represent a linear basis of the set of polynomials of degree d in the homogenized ideal $\langle F \rangle^h$, so G_h is a $d_{\max}$-Gröbner basis of $\langle F \rangle^h$, and G is a $d_{\max}$-Gröbner basis of $\langle F \rangle$.

For the last part of the proposition, note that the GRevLex order is a degree order, and so the leading term of f is the leading term of its highest degree component. Since there is no degree fall in the algorithm, all pivots used for computing the echelon form, and thus all leading terms of a polynomial in G_h, are monomials in $X_1, \dots, X_n$ (they are not divisible by H). This shows that $\mathsf{LT}(\langle F \rangle) = \mathsf{LT}(\langle F_h \rangle)$, and since generators of $\mathsf{LT}(\langle F_h \rangle)$ can be computed at degree d_{reg}, the same follows for $\langle F \rangle$. □

Remark 2.43. If F is a polynomial system which forms a regular sequence in the affine sense, then the homogenized ideal of $\langle F \rangle$ is generated by the homogenization of F:

$$\langle F \rangle^h = \langle F^h \rangle$$

This is false in general. For example, the system from Example 2.37 is equal to $\langle X, Y \rangle$, and so it is homogeneous. But $\langle f^h, g^h \rangle = \langle X, HY \rangle$.

In general, the relationship between the homogenized ideal and the homogenized generators involves a saturation by H:

$$\langle F \rangle^h = \langle F^h \rangle : H^\infty$$

2.4.4. Change of order: FGLM

In this section, we present algorithm FGLM ([Fau+93]), which allows for change of ordering in the zero-dimensional case. Informally, it converts a Gröbner basis for some monomial order into a Gröbner basis for another order.

The reason such algorithms were developed is because the direct algorithms such as F_5 are generally much more efficient at computing bases for some orders than others. Typically, the GRevLex is usually the easiest order, and the Lex order is the hardest. Elimination orders stand in between, depending on how many variables we wish to eliminate.

By using a change of order algorithm, we may first compute a Gröbner basis for an "easy" order and then use the change of order to compute the wanted basis.

Algorithm FGLM proceeds as follows: let I be a zero-dimensional of $\mathbb{K}[\mathbf{X}]$, let G_1 be a Gröbner basis of I for some monomial order $<_1$ and let $<_2$ be another order for which we want a Gröbner basis of I. Since I is zero-dimensional, $\mathbb{K}[\mathbf{X}]/I$ is a vector space of finite dimension $D = \deg(I)$. The algorithm will consider two monomial bases for this vector space, namely the two *staircases* for the monomial orders at hand.

Definition 2.44. Let I be a zero-dimensional ideal of $\mathbb{K}[\mathbf{X}]$, and let $<_{\text{mon}}$ be a monomial order. The *staircase* of I with respect to $<_{\text{mon}}$ is the set of monomials which are not leading monomial of a polynomial in I:

$$\mathcal{S}_I = \{m \in \mathcal{M} \mid m \notin \mathsf{LT}(I)\}$$

If G is a Gröbner basis of I with respect to $<_{\text{mon}}$, then

$$\mathcal{S}_I = \left\{m \in \mathcal{M} \mid \forall\, g \in G, \mathsf{LT}(g) \nmid m\right\}.$$

The main idea of Algorithm FGLM (Algorithm 2.5) is to use linear algebra to model the structure of the quotient ring:

1. Use G_1 to compute the staircase $\mathcal{S}_1$ of I with respect to $<_1$;
2. Compute the matrices of the multiplication by $x_1, \ldots, x_n$ in the basis $\mathcal{S}_1$;
3. Compute the staircase $\mathcal{S}_2$ of I with respect to $<_2$

The two first steps are performed by the subroutine MultiplicationMatrix (Algorithm 2.4): it is a routine taking as input a Gröbner basis G of an ideal I and returning the matrices of all linear maps $\mathbb{K}[\mathbf{X}]/I \longrightarrow \mathbb{K}[\mathbf{X}]/I$ given by the multiplication by X_i followed by the normal form in the monomial basis $\mathcal{S}$, for i ranging in $\{1, \ldots, n\}$. The number of normal forms to compute can be limited by noticing that any monomial $X_i\mu$, $\mu \in \mathcal{S}$ falls in one of the following cases:

- $X_i\mu \in \mathcal{S}$, in which case $\mathsf{NF}(X_i\mu) = X_i\mu$;
- $X_i\mu = \mathsf{LT}(g)$, $g \in G$, in which case $\mathsf{NF}(X_i\mu) = g - X\mu$;
- otherwise, if $\mathsf{NF}(\mu) = \sum_k \alpha_k\mu_k$, then $\mathsf{NF}(X_i\mu) = \sum_k \alpha_k \mathsf{NF}(X_i\mu_k)$.

In the two former cases, the normal form can be computed for free. In the latter case, it is given as a linear combination of normal forms of smaller monomials (according to the monomial order).

Algorithm 2.4 MultiplicationMatrix

Input: G a Gröbner basis of a zero-dimensional ideal I for a monomial order $<_{\text{mon}}$
Output: $\mathbf{M} = (M_1, \ldots, M_n)$ multiplication matrices of $X_1, \ldots, X_n$

$\mathcal{S} \leftarrow \{\mu \text{ monomial of } \mathbb{K}[\mathbf{X}] \mid \mu \notin \mathsf{LT}(\langle G \rangle)\}$
$F \leftarrow \{X_i\mu, i \in \{1, \ldots, n\}, \mu \in \mathcal{S}\} \setminus \mathcal{S}$
$N \leftarrow []$
for $\mu \in \mathcal{S}$ **do** $N[\mu] \leftarrow \mu$ **end for**
Sort F by increasing $<_{\text{mon}}$ order
for $\mu \in F$ **do**
    if $\mu \in \mathsf{LT}(G)$ **then**
        Find g such that $g \in G$ and $\mu = \mathsf{LT}(g)$
        $N[\mu] \leftarrow g - \mu$
    else
        Find (i, μ') such that $\mu = X_i\mu'$, $\mu' \in F$
        Write $N[\mu'] = \sum_{\mu'' <_{\mu'}} \alpha_{\mu''}\mu''$
        $N[\mu] \leftarrow \sum_{\mu'' <_{\mu'}} \alpha_{\mu''} N[X_i\mu'']$
    end if
end for
Copy the normal forms N into the multiplication matrices $\mathbf{M}$
return $\mathbf{M}$

Algorithm FGLM (Algorithm 2.5) then computes the staircase $\mathcal{S}_2$ of the ideal with respect to $<_2$ maintaining a association between the monomials μ and their normal form v_μ *w.r.t.* G_1.

For this purpose, it considers, in increasing $<_2$ order, all monomials already known to be in $\mathcal{S}_2$ (starting with 1) and considers their smallest multiples. Let μ be a monomial in the staircase, and let X_i be a variable. The normal form of $X_i\mu$ with respect to G_1 can be computed by matrix multiplication:

$$v_{X_i\mu} \coloneqq \mathsf{NF}(X_i\mu) = M_i \cdot v_\mu.$$

If this normal form is linearly independent from all previously computed normal forms, this means that $X_i\mu$ lies in the staircase $\mathcal{S}_2$. Otherwise, $X_i\mu$ lies in the initial ideal of I *w.r.t.* $<_2$, and the linear relation between $(v_s)_{s \in \mathcal{S}_2}$ and $v_{X_i\mu}$ gives a polynomial of I with leading term $X_i\mu$: this polynomial may be added to the Gröbner basis of I *w.r.t.* $<_2$.

Algorithm 2.5 FGLM ([Fau+93])

Input: G_1 a Gröbner basis of I for a monomial order $<_1$, another monomial order $<_2$
Output: G_2 Gröbner basis of I for $<_2$

if $G_1 = \{1\}$ **then return** $\{1\}$ **end if**
$\mathbf{M} = (M_1, \dots, M_n) \leftarrow \mathsf{MultiplicationMatrix}(G_1)$
$G_2 \leftarrow \varnothing$
$\mathcal{S}_2 \leftarrow \{1\}$
$v_1 \leftarrow (1, 0, \dots, 0)$
while $\exists\, i \in \{1, \dots, n\}, \mu \in \mathcal{S}_2, X_i\mu \notin \mathcal{S}_2$ **do**
    $(X_i, \mu) \leftarrow \min\{(X_i, \mu) \mid i \in \{1, \dots, n\}, \mu \in \mathcal{S}_2, X_i\mu \notin \mathcal{S}_2 \cup \mathsf{LT}_{<_2}(G_2)\}$
    $v_{X_i\mu} \leftarrow M_i \cdot v_\mu$
    if $v_{X_i\mu} \in \mathrm{Vect}\{v_s \mid s \in \mathcal{S}_2\}$ **then**
        Find $(\lambda_s)_{s\in\mathcal{S}_2}$ such that $v_{X_i\mu} = \sum_{s\in\mathcal{S}_2} \lambda_s v_s$
        Add $X_i\mu - \sum_{s\in\mathcal{S}_2} \lambda_s s$ to G_2
    else
        Add μ to $\mathcal{S}_2$
    end if
end while
return G_2

2.5. Complexity results

2.5.1. Complexity model and notations

In this section, we give complexity bounds for Algorithms Matrix-F_5 and FGLM. We measure the *arithmetic complexity* of algorithms, that is the number of operations (additions, multiplications, inversions) in the coefficient field $\mathbb{K}$.

This corresponds to time complexity if we assume that field operations are done in constant time. This is a reasonable assumption for finite fields. On the other hand, for infinite fields, it means that we ignore the growth of the coefficients of the polynomials.

Because of the central role of linear algebra in Gröbner basis algorithms, complexity estimates shall frequently depend on the constant ω defined as the smallest positive number such that one can compute the product of two $k \times k$ matrices in time $O(k^\omega)$.

Most operations on $k\times k$ matrices can be reduced to matrix multiplications, and thus performed in time $O(k^\omega)$: this includes the row echelon form, the characteristic polynomial...

The exact value of ω is unknown. The best known theoretical bound at the time of writing is $\omega < 2.372\,863\,9$ ([LeGall14]). In practice, matrix multiplication is typically done either

- using the naïve sequential implementation, with a complexity exponent of 3, or
- using Strassen's algorithm ([Str69]), with a complexity exponent of $\log_2(7) \approx 2.81$.

Finally, it is important to note that complexity estimates are always given under genericity assumptions. The worst-case complexity of computing a Gröbner basis is known to be doubly-

exponential in the degree of the polynomials ([MM82; BS88]), but this bound is not reached for generic systems.

2.5.2. Complexity of Matrix-F_5 and degree of regularity

In the case of Matrix-F_5, at each degree d, we build incrementally a matrix M_d whose columns are indexed by the monomials of degree d in $\mathbb{K}[\mathbf{X}]$. This matrix is then reduced to echelon form, and this step is the most expensive of the algorithm. Hence the complexity of the algorithm depends mainly on the size of the matrices being built, and on the number of such matrices. Both depend on the highest degree reached when running the algorithm.

Recall that if the system F forms a regular sequence, then the F_5 criterion eliminates all reductions to zero. As a consequence, the number of rows of the matrices is necessarily less than the number of columns.

As for the number of matrices that we need to compute and reduce, it is bounded by the stopping degree $d_{\max}$, given as an input to the algorithm. All in all, it gives the following simple estimate for the complexity of algorithm Matrix-F_5:

Theorem 2.45 ([BFS14, Prop. 1]). *Let $F \subset \mathbb{K}[\mathbf{X}]$ be a homogeneous polynomial system and $d_{\max} \in \mathbb{N}$, then Matrix-$F_5(F, d_{\max})$ computes a $d_{\max}$-Gröbner basis of $\langle F \rangle$ in time*

$$O\left(d_{\max}\binom{n + d_{\max} - 1}{d_{\max}}^3\right)$$

We now focus on a typical usage of algorithm F_5: computing a GRevLex basis for a system defined by a sufficiently generic system. In this case, we want to estimate how large $d_{\max}$ needs to be in order for Matrix-F_5 to return a Gröbner basis of the ideal.

Definition 2.46. Let F be a system of homogeneous polynomials. The *degree of regularity* of F is the smallest degree d_{reg} such that Matrix-$F_5(F, d_{\text{reg}}, <_{\text{grevlex}})$ gives a Gröbner basis of $\langle F \rangle$.

Recall that we defined the index of regularity of an ideal as the degree of its Hilbert series (Definition 1.32). This notion gives a bound for the degree of regularity of regular sequences, under genericity hypotheses:

Theorem 2.47 (Macaulay's bound). *Let $F = (f_1, \ldots, f_m)$ be a system of homogeneous polynomials with respective degree $(d_1, \ldots, d_m)$, and assume that F is in Noether position with respect to the variables $X_{m+1}, \ldots, X_n$. Then*

$$d_{\text{reg}}(F) \leq i_{reg}(\langle F \rangle) + 1 = \sum_{i=1}^{m} (d_i - 1) + 1.$$

Remark 2.48. Recall that if $m = n$, that is if the ideal $\langle F \rangle$ is zero-dimensional, the condition of being in Noether position is equivalent to F being a regular sequence.

Proof. First assume that $m = n$. Then the ideal $\langle F \rangle$ is zero-dimensional, and the Hilbert series of $\mathbb{K}[\mathbf{X}]/\langle F \rangle$ is a polynomial with degree i_{reg}. If $d \geq i_{\text{reg}} + 1$, all monomials in $\mathbb{K}[\mathbf{X}]$ are in $\langle F \rangle$, and

so any polynomial with degree greater than $i_{\text{reg}} + 1$ is reducible modulo some polynomial in $\langle F \rangle$ with lower degree.

Now assume that $m < n$. Since F is in Noether position, from Proposition 2.9, if G is a Gröbner basis of F, then $G \cup \{X_{m+1}, \ldots, X_n\}$ is a Gröbner basis of $\langle F, X_{m+1}, \ldots, X_n \rangle$. Since this is a zero-dimensional regular sequence, the bound above applies. □

Theorem 2.49 ([BFS14, Prop. 1]). *Let $F = (f_1, \ldots, f_m)$ be a system of homogeneous polynomials. Then a run of Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a $d_{\max}$-Gröbner basis of $\langle F \rangle$ in time*

$$O\left(d_{\max}\binom{n + d_{\max} - 1}{d_{\max}}^3\right). \tag{2.2}$$

If $d_{\max} \geq d_{\text{reg}}$ and the order is GRevLex, then Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a Gröbner basis of F. If F is in Noether position with respect to the variables $X_{m+1}, \ldots, X_n$, then d_{reg} is bounded by

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - 1) + 1.$$

In the affine case, for systems regular in the affine sense, Proposition 2.42 ensures that we can run AlgorithmMatrix-F_5 on the homogenized system up to the degree of regularity of the highest degree components, and obtain a Gröbner basis. So the same complexity bounds hold, using the number of monomials of degree at most d for the size of the matrix at degree d.

Theorem 2.50. *Let $F = (f_1, \ldots, f_m)$ be a system of polynomials which are regular in the affine sense, and assume that we are using the GRevLex order. Then Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a $d_{\max}$-Gröbner basis of $\langle F \rangle$ in time*

$$O\left(d_{\max}\binom{n + d_{\max}}{d_{\max}}^3\right).$$

If $d_{\max} \geq d_{\text{reg}}$, then Matrix-$\mathsf{F}_5 F, d_{\max}$ computes a Gröbner basis of F. If the highest degree components of F are in Noether position with respect to the variables $X_{m+1}, \ldots, X_n$, then d_{reg} is bounded by

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - 1) + 1.$$

2.5.3. Thin-grained complexity of Matrix-F_5

The complexity bound in the previous section is far from optimal: it is obtained by taking the complexity of reducing one dense matrix as large as the largest one the algorithm sees, and multiplying this complexity by the number of matrices that the algorithm needs to reduce.

There are two ways this bound can be improved: by taking into account the fact that the matrices are not dense, and by looking at the size of all matrices instead of the largest one.

Both refinements come as a consequence of the following *structure lemma*:

Lemma 2.51 (Structure lemma, [BFS14, Prop. 11]). *Let $F = (f_1, \dots, f_m)$ be a homogeneous system of polynomials of respective degree $(d_1, \dots, d_m)$, in simultaneous Noether position with respect to the order $X_1 > \cdots > X_m$. For all $i \in \{1, \dots, m\}$, let G_i be the reduced GRevLex Gröbner basis of $\langle f_1, \dots, f_i \rangle$. Let $i \in \{1, \dots, m\}$, and $g \in G_i$ with signature (j, μ). Then:*

- $j \leq i$
- $\mathsf{LT}(g) \in A_i := [X_1, \dots, X_i]$
- $\mu \in A_{i-1} = \mathbb{K}[X_1, \dots, X_{i-1}]$

Proof. Since G_i is a Gröbner basis of $\langle f_1, \dots, f_i \rangle$, any polynomial in G_i can be written as an algebraic combination of these polynomials, and thus $j \leq i$.

For the second statement, without loss of generality, we may assume that $i = j$. Indeed, if g has signature (j, μ) with $j < i$, then $g \in G_j$ and applying the proposition to G_j yields that $\mathsf{LT}(g) \in \mathbb{K}[X_1, \dots, X_j] \subset \mathbb{K}[X_1, \dots, X_i]$. From Corollary 2.9, $\mathsf{LT}(\langle f_1, \dots, f_i \rangle)$ is generated by monomials in A_i, and since $g \in G_i$, $\mathsf{LT}(g) \in A_i$.

In order to prove the third statement, we proceed by induction on the signature (i, μ). For $i = 1$, there is no reduction so the result holds. Let us assume $i > 1$, then the lemma is true for the minimal valid signature $(i, 1)$. Let now f be a polynomial appearing in the computation of G_i, in W-degree $d > d_i$. Before any reduction, f can be written as $X_l f_0$, where f_0, with signature (i, μ) is reduced against the basis G_{i-1}, and where X_l is a variable with $l \leq i$.

If there is a polynomial in G_i with signature (i, μ), it means that the row corresponding to f was reducible (in the linear sense) by the rows above it. Algebraically, it means that there exists g in G_k, $k \leq i$, having signature $(k, \beta) < (i, \mu)$, such that f is head-reducible modulo g. By induction hypothesis, we find that $\beta \in A_{k-1}$, and from the second statement, we deduce that $\mathsf{LT}(g) \in A_k$. Since g head-reduces f, $\mathsf{LT}(g)$ divides $\mathsf{LT}(f) = X_l \mathsf{LT}(f_0)$, and since f_0 is reduced, X_l divides $\mathsf{LT}(g)$, and so $X_l \in A_k$. Now we are in one of the three following cases:

1. $k < i$: then $X_l \in A_{i-1}$, and the label of f is $(i, X_l \mu)$ with $X_l \mu \in A_{i-1}$;
2. $k = i$, $l < i$: same;
3. $k = i$, $l = i$: then we can write $X_i \mathsf{LT}(f_0) = c\mathsf{LT}(g)$, with $c \in A_i$. If $X_i \mid c$, $\mathsf{LT}(g)$ divides $\frac{X_i}{X_i}\mathsf{LT}(f_0)$, which is absurd because we assumed f_0 to be reduced. So we must have $c \in A_{i-1}$. Then the label of cg is $(i, c\beta)$, which is greater than the label of f: $c\beta$ is in A_{i-1}, unlike $X_i \mu$. But that is absurd, because we assumed cg to reduce f. □

The structure lemma gives a correspondence between signatures and polynomials in the Gröbner basis. Counting the possible signatures thus gives us a bound on the number of polynomials in the Gröbner basis:

Proposition 2.52 ([BFS14, Th. 12]). *With the same notations, let $i \in \{1, \dots, m\}$ and $d \in \mathbb{N}$. The number of polynomials of degree d in G_i whose leading term does not belong to $\mathsf{LT}(G_{i-1})$ is bounded by the generic term $b_d^{(i)}$ of the following generating series:*

$$B_d^{(i)}(T) := \sum_{d=0}^{\infty} b_d^{(i)} T^d := T^{d_i} \prod_{k=1}^{i-1} \frac{1 - T^{d_k}}{1 - T}.$$

Remark 2.53. This proposition bounds the number of new elements added to the Gröbner basis by each polynomial f_i. It is tempting to write it as $G_i \setminus G_{i-1}$, but elements of G_{i-1} may be tail-reducible modulo f_i, changing their value in G_i. Another way of stating this proposition is that we count the number of polynomials in G_i with signature $(i, \bullet)$.

Proof. By Lemma 2.51, any polynomial in G_i with signature (i, μ) is such that $\mu \in A_{i-1} = \mathbb{K}[X_1, \dots, X_{i-1}]$. Furthermore, the F_5 criterion ensures that $\mu \notin F_{i-1} := \langle f_1, \dots, f_{i-1} \rangle$. The number of monomials of degree d in A_{i-1} such that $\mu \notin \langle f_1, \dots, f_{i-1} \rangle$ is precisely the dimension of $(A_{i-1}/\langle \theta_{i-1}(F_{i-1}) \rangle)_d$ as a $\mathbb{K}$-vector space. In other words, the number of available monomials for a signature of degree d is bounded by the generic term of the Hilbert series of $A_{i-1}/\langle \theta_{i-1}(F_{i-1}) \rangle$. Since F is in simultaneous Noether position, by Proposition 1.44, $\theta_{i-1}(F_{i-1})$ is a regular sequence, and this Hilbert series is:

$$\mathrm{HS}_{A_{i-1}/\langle \theta_{i-1}(F_{i-1}) \rangle} = \frac{\prod_{k=1}^{i-1} 1 - T^{d_k}}{(1-T)^{i-1}}$$

and the result follows, shifting this series by T^{d_i} to count the number of polynomials that we can form with these signatures. □

Theorem 2.54 ([BFS14, Sec. 3.2, Eq. 10])**.** *Let $F = (f_1, \dots, f_m)$ be a system of homogeneous polynomials with respective degrees $d_1, \dots, d_m$, and such that F is in simultaneous Noether position with respect to the order $X_1 > \cdots > X_n$.*

Then a run of Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a $d_{\max}$-Gröbner basis of $\langle F \rangle$ in time

$$O\left(\sum_{i=1}^{m} \sum_{d=0}^{d_{\mathrm{reg}}} b_d^{(i)} \binom{i+d-1}{d} \binom{n+d-1}{d} \right). \tag{2.3}$$

Proof. For each polynomial f_i, at each degree d, we reduce at most $b_d^{(i)}$ rows in the matrix. Each reduction uses at most $\binom{i+d-1}{d}$ pivots (because the resulting leading term is in $\mathbb{K}[X_1, \dots, X_i]$), and each elementary reduction costs at most $\binom{n+d-1}{d}$ arithmetic operations (the number of columns in the matrix). □

Both bounds are plotted for comparison in Figure 2.1.

2.5.4. Complexity of FGLM

As before, let I be a zero-dimensional ideal and let D be its degree.

Algorithm FGLM has two main steps: computing the multiplication matrices, and computing the Gröbner basis with respect to the new monomial order.

The cost of computing the multiplication matrices with Algorithm 2.4 is concentrated in the computation of the normal forms of the monomials of the form $X_i \mu'$, $\mu' \in F$. Each of these normal forms is computed as the product of a $K \times K$ matrix and a $1 \times K$ vector, where $K \leq D$.We need to compute at most $\#F \leq nD$ such normal forms, and overall, computing the multiplication matrices can be done in time

$$O(nD \times D^2) = O(nD^3).$$

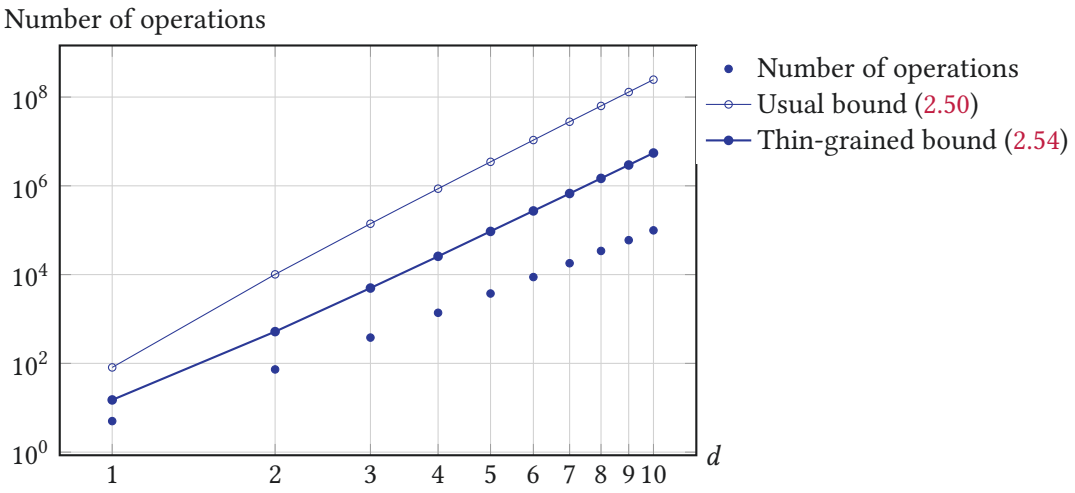

Figure 2.1: Number of operations in a run of algorithm Matrix-F_5 on a system of 3 generic homogeneous polynomials with degree d

Computing the Gröbner basis is done by incrementally building a linear basis of the staircase of I with respect to the second order. If we maintain a matrix of the basis throughout the algorithm, each step of verification of linear independence and recovery of the coefficients λ_j can be done by reducing one vector by this matrix. Overall, it is equivalent to computing the normal form of the whole matrix using the Gauss algorithm. This matrix has D columns and at most nD rows, so the complexity is again $O(nD^3)$.

Theorem 2.55 ([Fau+93]). *The complexity of algorithm FGLM for computing the Gröbner basis of a zero-dimensional ideal with degree D is*

$$O(n\,D^3).$$

Faster variants of FGLM have been proposed. In [FM13], the authors proposed algorithms taking advantage of the sparsity of the multiplication matrices. In particular, for a lexicographical basis in shape position, it suffices to compute the minimal polynomial of M_n, and this can be done in time $O(nD^\omega)$ using fast linear algebra. In [Fau+14], the authors show that, up to a generic change of coordinates, the multiplication matrices can be read at no cost on a GRevLex basis of the ideal. Overall, up to a generic change of coordinates, computing a Lex Gröbner basis from a GRevLex one can be done in time

$$O(nD^\omega).$$

Part II

Contributions

Chapter 3

Weighted homogeneous systems

The results presented in this chapter are extracted from a joint work with Jean-Charles Faugère and Mohab Safey El Din. They have been published in [FSV13] and [FSV15].

In this chapter, we consider weighted-homogeneous systems, as defined in Section 1.1.3. We define some properties of these systems, and we rework some characterizations which were available for homogeneous systems. We exhibit a computational strategy making use of existing algorithms, and show, under genericity assumptions, that the complexity is divided by $(\prod w_i)^\omega$ when compared with homogeneous systems with the same degree. Finally, we provide some experimental results showing that taking advantage of the weighted homogeneous structure can yield substantial speed-ups.

3.1. Introduction

Consider a vector of positive integers $W = (w_1, \ldots, w_n)$. As in Section 1.1.3, we define a gradation on $\mathbb{K}[X_1, \ldots, X_n]$ by defining the W-degree of a monomial as:

$$\deg_W(X_1^{\alpha_1} \cdots X_n^{\alpha_n}) = w_1\alpha_1 + \cdots + w_n\alpha_n.$$

An equivalent definition is that a polynomial f is W-homogeneous with W-degree d if and only if $f(X_1^{w_1}, \ldots, X_n^{w_n})$ is homogeneous with degree d.

There are two main strategies for computing Gröbner bases for weighted homogeneous systems:

- apply existing algorithms, disregarding the weighted homogeneous structure;
- use the above property to transform a weighted homogeneous system and apply existing algorithms to this system, exploiting its new homogeneous structure.

Experimentally, the second strategy appears to be faster than the first one for computing a GRevLex basis using F_4 or F_5. But when needed, the change of order using FGLM becomes a major bottleneck.

To the best of our knowledge, there is no way to evaluate the complexity of the first strategy, and complexity analyses for the second strategy only rely on the homogeneous structure of the system.

In this work, we justify that the second strategy is correct, and is actually a way to efficiently take advantage of the weighted homogeneous structure with pairs-based algorithms, such as F_4 or F_5. We also explain the bottleneck in the FGLM algorithm for zero-dimensional systems, and propose a computational strategy working around this obstacle, effectively dividing the complexity of FGLM by a factor $(\prod w_i)^3$.

We also give complexity results for the F_5 algorithm. For this purpose, we adapt algorithm Matrix-F_5 to the weighted case, and show that the bounds of Proposition 1.20 on the number of monomials in a W-graded algebra allow to divide the complexity bounds by $(\prod w_i)^3$. We also show that the complexity in practice is improved, because the degree of regularity of a weighted homogeneous system is lower than that of a homogeneous system with the same degrees.

As in the homogeneous case, complexity studies are obtained under some generic assumptions. More precisely, we give bounds on the degree of regularity for:

- zero-dimensional systems defined by a regular sequence;
- positive-dimensional systems defined by a sequence in Noether position with respect to the smallest variables;
- zero- or positive-dimensional systems defined by a sequence in simultaneous Noether position;

In the over-determined case, we define semi-regular weighted homogeneous sequences, and we show that under some hypotheses on the weights and the degrees, this property admits the same characterization as in the homogeneous case. This can be used to bound the degree of regularity of weighted semi-regular sequences. We give an example of such computations, by transposing an asymptotic analysis [Bar+05] of the degree of regularity of semi-regular sequences with $m = n + k$ to the weighted homogeneous case.

Prior work The special case $W = (1, \ldots, 1)$ is the usual homogeneous case. In this case, all properties and characterizations shown in this chapter specialize to known results that we described in Chapters 1 and 2.

Weighted homogeneous systems have been studied before, from the angle of singularity theory and commutative algebra. In particular, the results about the Hilbert series and the Hilbert function of weighted homogeneous ideals (see Chapter 1) can be found in most commutative algebra textbooks.

The computational strategy for systems with a weighted structure is not new either. For example, it is already implemented (partially: for weighted homogeneous or heuristically, systems containing some weighted homogeneous polynomials) in the computer algebra system Magma [Magma]. Additionally, the authors of [Tra96] proposed another way of taking into account the weighted structure, using the Hilbert series of the ideal. The authors of [CDR96] generalized this algorithm to systems homogeneous with respect to a multigraduation. Their definition of a system of weights is more general than the one we use in this thesis.

To the best of our knowledge, nobody presented a formal description of a computational strategy for systems with a weighted homogeneous structure (not necessarily weighted homogeneous), together with complexity estimates.

Some of the results presented in this chapter appeared in a conference paper [FSV13]: these results are the weak form of the weighted Macaulay bound (3.2) and the formal description of the algorithmic strategy for weighted homogeneous systems, with the complexity estimates (3.1) and (3.4). The remaining parts of this chapter were presented in the journal article [FSV15]: the conference paper lacked a hypothesis (reverse chain-divisible systems of weights), and therefore lacked the precise description of Hilbert series required to obtain results for semi-regular sequences. The sharp variant of the weighted Macaulay bound (3.3), under the assumption of simultaneous Noether position, was also added in the extended paper. Finally, the benchmarks section of the journal paper contains additional systems, arising in polynomial inversion problems.

The conference paper was using *quasi-homogeneous* instead of *weighted homogeneous* to describe the studied structure. While both names exist in the literature, *weighted homogeneous* seems to be more common, and to better convey the notion that this structure is a generalization of homogenity, instead of an approximation. The same notion is sometimes also named simply *homogeneous* (in which case the weights are determined by the degree of the generators; see for example [Eis95]), or homogeneous for a *nonstandard graduation* [DS06].

Main results By definition, weighted homogeneous polynomials can be made homogeneous by raising all variables to their weight. The resulting system can then be solved using algorithms for homogeneous systems. However, it appears experimentally that solving such systems is much faster than solving generic homogeneous systems. In this chapter, we show that the complexity estimates for homogeneous systems can be divided by $(\prod w_i)^\omega$ in case the system was originally W-homogeneous, ω being the complexity exponent of linear algebra operations ($\omega = 3$ for naive algorithms, such as the Gauss algorithm).

These complexity estimates depend on two parameters of the system: its *degree of regularity* d_{reg} and its *degree* $\deg(I)$. Both parameters can be obtained from the *Hilbert series* of the ideal, which can be precisely described under generic assumptions. To be more specific, we will consider systems defined by a *regular sequence* and systems which are in *simultaneous Noether position.*

Theorem. *Let $W = (w_1, \ldots, w_n)$ be a system of weights, and $F = (f_1, \ldots, f_m)$ a zero-dimensional W-homogeneous system of polynomials in $\mathbb{K}[X_1, \ldots, X_n]$, with respective W-degree $d_1, \ldots, d_m$. The complexity (in terms of arithmetic operations in $\mathbb{K}$) of Algorithm F_5 to compute a W-GRevLex Gröbner basis of $I := \langle F \rangle$ is bounded by*

$$C_{\mathrm{F}_5} = O\left(\frac{1}{(\prod w_i)^3} \cdot \binom{n + d_{\text{reg}} - 1}{d_{\text{reg}}}^3\right). \tag{3.1}$$

If F is a regular sequence (and in particular $m = n$), then d_{reg} can be bounded by the weighted Macaulay bound*:*

$$d_{\text{reg}} \leq \sum_{i=1}^{n} (d_i - w_i) + \max\{w_j\}. \tag{3.2}$$

If additionally F is in simultaneous Noether position w.r.t the order $X_1 > \cdots > X_n$, then the weighted Macaulay bound can be refined:

$$d_{\text{reg}} \leq \sum_{i=1}^{n} (d_i - w_i) + w_n. \tag{3.3}$$

The complexity of Algorithm FGLM to perform a change of ordering is bounded by

$$C_{FGLM} = O\left(n(\deg(I))^{\omega}\right). \tag{3.4}$$

Recall that if F forms a regular sequence, then $\deg(I)$ *is given by the* weighted Bézout bound

$$\deg(I) = \frac{\prod_{i=1}^{n} d_i}{\prod_{i=1}^{n} w_i}.$$

In particular, the bound (3.3) indicates that in order to compute a Gröbner basis faster for a generic enough system, one should order the variables by decreasing weights whenever possible.

The hypotheses of the theorem are not too restrictive. In the homogeneous case, regularity and simultaneous Noether position are generic properties. However, in the weighted homogeneous case, there are systems of weights and systems of weighted degrees for which they are not generic. We shall identify large families of systems of weights and systems of weighted degrees for which they are (Proposition 3.19).

All sequences in simultaneous Noether position are regular. Conversely, in the homogeneous case, all regular sequences are in simultaneous Noether position up to a generic linear change of coordinates. In the weighted homogeneous case, this is no longer true. In fact, there are even systems of weights for which there exists no non-trivial change of coordinates.

To avoid this limitation, we consider *reverse chain-divisible* systems of weights, that is systems of weights such that $w_n \mid w_{n-1} \mid \ldots \mid w_1$. This property ensures that there are non-trivial changes of coordinates of the form $X_i \leftarrow X_i + P_i(X_{i+1}, \ldots, X_n)$ for all i, with P_i a W-homogeneous polynomial with W-degree w_i. Under this assumption, many properties from the homogeneous case remain valid in a weighted setting. In particular, any regular sequence is in simultaneous Noether position up to a W-homogeneous change of coordinates (Theorem 3.15).

For many systems from practical applications, the weights can be chosen to be reverse chain-divisible. We give a few examples in the last section of this chapter.

If $m > n$, there is no regular sequence. Instead, we will consider systems defined by a *semi-regular* sequence, that is systems for which no reduction to zero appear in a run of Algorithm F_5. This property has several equivalent definitions in the homogeneous case. While these definitions can be easily extended to the weighted case, their equivalence is not necessarily true. However, we prove that these definitions are equivalent in the special case where the weights form a reverse chain-divisible sequence.

In the homogeneous case, the property of being semi-regular is only conjectured to be generic, but this conjecture is proved in a handful of cases [Mor96, Thm. 1.5]. In this chapter, we adapt the proof of one of these cases, namely the case $m = n + 1$ in a base field of characteristic 0.

For semi-regular systems with $m = n + 1$, we obtain a bound on the degree of regularity. More generally, in the homogeneous case, one can compute asymptotic estimates on the degree

of regularity of a semi-regular sequence [Bar+05; Bar04]. These estimates can be adapted to the weighted homogeneous case. As an example, we give an asymptotic bound on the degree of regularity for semi-regular systems with $m = n + k$ for a given integer k:

Theorem. *Let n and k be two positive integers, and let $m = n + k$. Let w_0 and d_0 be two positive integers such that $w_0 \mid d_0$. Consider the system of n weights $W = (w_0, \ldots, w_0, 1)$. Let F be a semi-regular sequence in $\mathbb{K}[X_1, \ldots, X_n]$, made of W-homogeneous polynomials with W-degree d_0. Then the highest degree reached in the computation of a W-GRevLex Gröbner basis of $\langle F \rangle$ is asymptotically bounded by*

$$d_{\text{reg}} = n \frac{d_0 - w_0}{2} - \alpha_k \sqrt{n \frac{d_0^2 - w_0^2}{6}} + O\left(n^{1/4}\right).$$

where α_k is the largest root of the k'th Hermite polynomial.

Experimentally, if we lift the assumption that the system of weights is reverse chain-divisible, the degree of regularity does not appear to rise too far beyond the bound. Future work on the topic could include characterizing the Hilbert series of W-homogeneous semi-regular sequences in full generality, in order to obtain bounds on the W-degree of regularity.

In practice, taking advantage of the weighted structure when applicable yields significant speed-ups. Some instance of a weighted structure has already been successfully exploited for an application in cryptography [Fau+13]. We also present timings obtained with several polynomial inversion problems, with speed-ups ranging from 1–2 to almost 100. In particular, we use these techniques in order to compute the relations between fundamental invariants of several groups [Stu08]. For some groups such as the Cyclic-5 group or the dihedral group D_5, computing these relations is intractable without considering the weighted structure of the system, while it takes only a few seconds or minutes when the weighted structure is exploited. All these systems are examples of applications where the weights giving the appropriate W-homogeneous structure are naturally reverse chain-divisible. These experimentations have been carried using F_5 and FGLM with the Gröbner basis library FGb [Fau10] and F_4 with the computer algebra system Magma [Magma].

There are other applications where Gröbner bases are computed for polynomial systems with a weighted-homogeneous structure, for example in coding theory, both for generating codes [BP99, sec. 5], [Leo09] and for decoding through Guruswami-Sudan's algorithm (see [GR09] for an overview).

Organisation of the chapter In Section 3.2, we describe some algebraic properties of W-homogeneous systems. In particular, we show that regular sequences and sequences in Noether position are generic, and we give a characterization of the Hilbert series of a W-homogeneous regular sequence. In Section 3.3, we give bounds for the degree of regularity of a generic zero-dimensional system, and in Section 3.3.4, we give some bounds for positive-dimensional systems as well. In Section 3.4, we consider semi-regular systems. We give some equivalent definitions of this property, and we show how asymptotic estimates of the degree of regularity can be adapted from the homogeneous case to the weighted case. Additionally, we prove that Fröberg's conjecture in the case $m = n + 1$ is true in the weighted case, as in the homogeneous case,

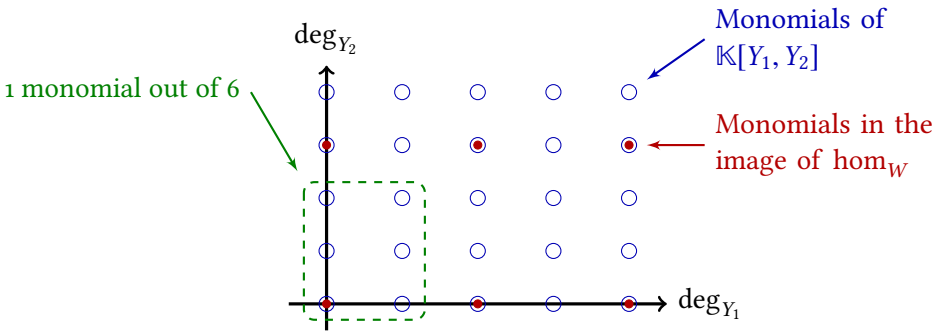

Figure 3.1: Counting monomials in the image of $\hom_W$ (here $W = (2, 3)$)

provided that the base field is large enough. In Section 3.5, we describe strategies for computing Gröbner bases for weighted homogeneous systems, and we give complexity estimates for these strategies. Finally, in Section 3.6, we show how weighted structures can appear in applications, and we give some benchmarks for each example.

3.2. Properties

3.2.1. Definitions and properties

As described in the introduction, one can make a W-homogeneous polynomial homogeneous using the following proposition.

Proposition 3.1. *Let $(\mathbb{K}[X_1, \ldots, X_n], W)$ be a graded polynomial algebra. Then the application*

$$\begin{array}{rccc} \hom_W : & (\mathbb{K}[X_1, \ldots, X_n], W) & \longrightarrow & (\mathbb{K}[Y_1, \ldots, Y_n], \mathbf{1}) \\ & f & \longmapsto & f(Y_1^{w_1}, \ldots, Y_n^{w_n}) \end{array}$$

is an injective graded morphism, and in particular the image of a weighted homogeneous polynomial is a homogeneous polynomial.

Remark 3.2. This morphism is *not* surjective. More precisely, the bounds and asymptotics for Sylvester denumerants (Proposition 1.20 and 1.19) apply, and show that on average, there are $\prod w_i$ times more monomials in $\mathbb{K}[\mathbf{X}]$ at degree d than at W-degree d.

The morphism $\hom_W$ gives an intuition for these asymptotic estimates (Figure 3.1). For example, consider $\mu = Y_1^{w_1-1} \cdots Y_n^{w_n-1}$. There is only 1 monomial in the divisors of μ which also lies in the image of $\hom_W$ (namely 1); there are $\prod w_i$ monomials dividing μ; and all non-trivial multiples of μ are divisible by more monomials in the image of $\hom_W$.

The W-GRevLex monomial ordering is the pullback of the GRevLex monomial ordering by this morphism:

$$u <_{W\text{-grevlex}} v \iff \hom_W(u) <_{\text{grevlex}} \hom_W(v).$$

Given a W-homogeneous system F, one can build the homogeneous system $\hom_W(F)$, and then apply classic algorithms (Section 2.4) to that system to compute a GRevLex or Lex Gröbner basis of the ideal generated by $\hom_W(F)$.

Definition 3.3. The *W-degree of regularity* of the system F is the highest degree $d_{\mathrm{reg},W}(F)$ reached in a run of F_5 to compute a GRevLex Gröbner basis of $\hom_W(F)$. When the graduation is clear in the context, we may call it degree of regularity, and denote it d_{reg}.

Remark 3.4. Unlike what we could observe in the homogeneous case, this definition depends on the order of the variables (we shall give an example in Table 3.1 in Section 3.3.2, and another, with timings, in Table 3.4 in Section 3.6.1).

As before, we only consider the *affine* varieties associated with the ideals we consider. In particular, for polynomials in $\mathbb{K}[X_1,\ldots,X_n]$, the dimension of $V(0)$ is n, and a zero-dimensional variety is defined by at least n polynomials if the base field is algebraically closed.

3.2.2. Degree and Bézout's bound

The degree of a weighted homogeneous ideal can be defined in the same way as the homogeneous case, using the Hilbert series:

Definition 3.5. Let I be a weighted homogeneous ideal with dimension d. Its *degree* is $\deg(I) = Q(1)$ where $Q = \mathrm{HS}_{A/I}(T)\cdot(1-T)^d$.

Remark 3.6. With this definition, the degree has a geometrical interpretation in the zero-dimensional case. Then $d = 0$, so $\deg(I) = \mathrm{HS}_{A/I}(1)$ is an integer, which corresponds to the dimension of A/I as a $\mathbb{K}$-vector space, that is the cardinality of the staircase of the ideal, or the number of solutions counted with multiplicity.

Remark 3.7. This is the definition used by the software Macaulay2 for the degree of a positive-dimensional variety (function `degree(Module)` in [Macaulay2]).

Remark 3.8. In the positive-dimensional case, even if the sequence F is in Noether position *w.r.t.* the variables $X_1,\ldots,X_m$, the degree of $\langle F\rangle$ is not necessarily the number of solutions of $F_{\mathrm{ext}} = (f_1,\ldots,f_m,X_{m+1},\ldots,X_n) = 0$. For example, consider $\mathbb{K}[X,Y]$ with the weights $W = (2,1)$, and let $f = X + Y^2$. The ideal $\langle f\rangle$ has degree 2, and indeed, a generic hypersurface of W-degree 2 is another parabola with horizontal axis, so that there are 2 intersection points (see Figure 3.2). However, (f) is in Noether position with respect to the variable Y, and $\langle f, Y\rangle = \langle X, Y\rangle$ has only 1 zero.

Remark 3.9. It may even happen that positive-dimensional varieties have a rational degree. For example, consider the ideal $\langle XY\rangle$ in $\mathbb{K}[X,Y]$ graded with respect to $W = (2,3)$. Then

$$\mathrm{HS}_{A/I}(T) = \frac{(1-T^5)}{(1-T^2)(1-T^3)} = \frac{1+T+T^2+T^3+T^4}{(1-T)(1+T)(1+T+T^2)}$$

and

$$\deg(I) = \frac{5}{6}.$$

Intuitively, computing the degree of a positive-dimensional variety can be done by cutting it by generic hypersurfaces. In the total degree case, one can use hyperplanes, that is degree 1 hypersurfaces, so that cutting does not change the degree. However, with $W = (2,3)$, the

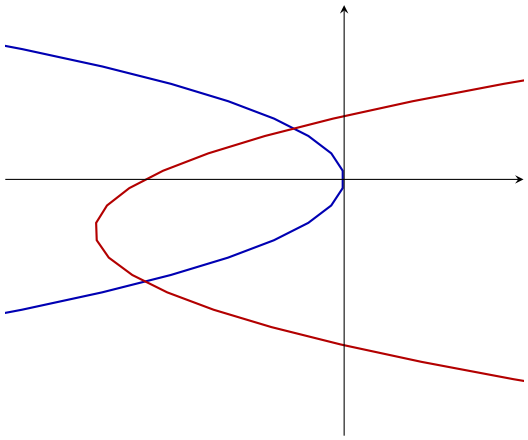

Figure 3.2: Common zeroes of $X^2 + Y$ (in blue) and a generic polynomial of $[1, 2]$-degree 2 (in red)

first W-degree at which one can really form *generic* hypersurfaces is 6, and intersecting XY with a degree 6 hypersurface generically yields 5 solutions: going through $\hom_W$, a generic intersection will have 30 points, and $\hom_W$ multiplies this number of solutions by 6.

The following proposition formalizes the intuitive notion that "*going through* $\hom_W$ *multiplies the number of solutions by the product of the weights*":

Proposition 3.10. *Let $W = (w_1, \dots, w_n)$ be a system of weights, and $F = (f_1, \dots, f_m)$ a sequence of W-homogeneous polynomials, with respective W-degrees $d_1, \dots, d_m$. Then*

$$\deg(F) = \frac{\deg(\hom_W(F))}{\prod_{i=1}^n w_i}.$$

Proof. Let $I = \langle F \rangle$ and $\hom_W(I) = \langle \hom_W(F) \rangle$. Write the Hilbert series of A/I (as an algebra graded in W-degree) and $A/\hom_W(I)$ (as an algebra graded in total degree) using Proposition 1.30:

$$\mathrm{HS}_{A/I}(T) = \frac{P(T)}{\prod_{i=1}^n (1 - T^w)}$$
$$\mathrm{HS}_{A/\hom_W(I)} = \frac{Q(T)}{(1-T)^n}.$$

The proof of Proposition 1.38 actually shows that $P = Q$, so

$$\mathrm{HS}_{A/I}(T) = \frac{\mathrm{HS}_{A/\hom_W(I)}(T)}{\prod_{i=1}^n (1 + T + \cdots + T^{w_{i-1}})},$$

and evaluating this equation at 1 yields the wanted result. □

A classic consequence of this is the weighted version of Bézout's bound:

Theorem 3.11. *Let $W = (w_1, \dots, w_n)$ be a system of weights, and $F = (f_1, \dots, f_m)$ a regular sequence of W-homogeneous polynomials, with respective W-degrees $(d_1, \dots, d_m)$. Then*

$$\deg(I) = \mathrm{HS}_{\mathbb{K}[X]/I}(1) = \frac{\prod_{i=1}^{m} d_i}{\prod_{i=1}^{n} w_i}.$$

3.2.3. Changes of variables and reverse chain-divisible systems of weights

Several properties from the homogeneous case turn out to be no longer true in the weighted case.

For example, many properties of polynomial systems are true up to a linear change of coordinates. Since linear changes of coordinates leave the **1**-homogeneous components of the polynomial algebra stable, it is also true of **1**-homogeneous polynomial systems. A good example of such a property is Noether's normalization theorem.

However, in a weighted setting, linear transformations on the variables do not necessarily preserve the W-degree of the monomials.

If we add some constraints on the system of weights, we may define non-trivial changes of variables. More precisely, we will consider *reverse chain-divisible* systems of weights, defined as follows.

Definition 3.12. A system of weights, W is *reverse chain-divisible* if

$$w_n \mid w_{n-1} \mid \dots \mid w_1$$

In this situation, the weights are coprime if and only if $w_n = 1$.

Remark 3.13. The name "chain-divisible" can be found in [Alf05], referring to a notion introduced in [Alf98].

In this setting, many results from the homogeneous case can now be adapted to the weighted homogeneous case. For example, Noether's normalization theorem (Theorem 1.88) is true in a weighted setting if we replace linear changes of coordinates by W-homogeneous changes of coordinates with degree (w_i).

Lemma 3.14 (Noether normalization lemma, weighted case)**.** *Let $\mathbb{K}$ be an infinite field, W be a reverse chain-divisible system of weights and $f \in R = \mathbb{K}[X_1, \dots, X_r]$ be a non-constant polynomial, W-homogeneous with W-degree d. Then there are elements $X'_1, \dots, X'_{r-1} \in R$ such that R is a finitely generated module over $\mathbb{K}[X'_1, \dots, X'_{r-1}, f]$. Furthermore, if the field has characteristic 0 or large enough, there exists a dense Zariski-open subset $U \subset \mathbb{K}^{r-1}$ such that for all $(a_1, \dots, a_{r-1}) \in U$, one can choose $X'_i = X_i - a_i X_r^{w_i/w_r}$.*

Proof. We follow the proof of Lemma 1.89. For any $1 \leq i \leq r-1$, let $a_i \in \mathbb{K}$, and let $X'_i = X_i - a_i X_r^{w_i/w_r}$. We need to show that for generic a_i, under this change of variables, f is monic in X_r. Under this change of variables, collecting the coefficients of f in X_r yields:

$$\begin{aligned} f(X_1, \dots, X_r) &= f(X'_1 + a_1 X_r^{w_1/w_r}, X'_2 + a_2 X_r^{w_2/w_r}, \dots, X_{r-1} + a_{r-1} X_r^{w_{r-1}/w_r}) \\ &= f(a_1, \dots, a_{r-1}, 1) X_r^d + \cdots \end{aligned}$$

So the set of all a_i's such that f is monic in X_r is exactly the set of all a_i's such that

$$f(a_1, \dots, a_{r-1}, 1) \neq 0,$$

and since f is W-homogeneous non-constant, this is a non-empty open subset of $\mathbb{K}^{r-1}$. □

Apart from the lemma, the proof of Noether's normalization theorem does not depend on the graduation, and as a consequence:

Theorem 3.15 (Noether's normalization theorem, weighted case)**.** *Let W be a reverse chain-divisible system of weights, and let F be a W-homogeneous zero-dimensional regular sequence in $\mathbb{K}[X_1, \dots, X_n]$. Then, for a generic choice of W-homogeneous polynomials P_i with W-degree w_i, the change of variable*

$$X_i = X'_i + P_i(X_{i+1}, \dots, X_n),$$

is such that $F(X_1(\mathbf{X}'), \dots, X_n(\mathbf{X}'))$ is in simultaneous Noether position with respect to the order $X'_1 > X'_2 > \cdots > X'_n$.

A central property of reverse chain-divisible weights is the following proposition. In the homogeneous case, if $d_1 \leq d_2$ are two non-negative integers, then any monomial with degree d_2 is divisible by a monomial with degree d_1. When the system of weights is reverse chain-divisible, the following proposition states a similar result for the weighted case.

Proposition 3.16. *Assume that $W = (w_1, \dots, w_n)$ is a system of weights, such that $w_1 \geq w_2 \geq \cdots \geq w_n$. The following statements are equivalent:*

1. *The system of weights W is reverse chain-divisible;*
2. *Let $d_1 \leq d_2$ positive integers, $i \in \{1, \dots, n\}$, and m_2 a monomial of W-degree d_2. Assume that w_i divides d_1, and that m_2 is not divisible by any of the variables $X_1, \dots, X_{i-1}$. Then there exists a monomial m_1 with W-degree d_1 such that $m_1 \mid m_2$.*

Proof. (1 $\implies$ 2). Fix d_1. We shall prove by induction over d_2 that for any monomial m_2 with W-degree d_2 satisfying the hypotheses of (2), there exists a monomial m_1 with W-degree d_1 dividing m_2. The case $d_2 = d_1$ is immediate, since we can use $m_1 \coloneqq m_2$.

Assume that $d_2 > d_1$, and let m_2 be a monomial of W-degree d_2. Let j be the greatest index of a variable dividing m_2, write $m_2 = X_j^\alpha m'_2$, where m'_2 is a monomial in $\mathbb{K}[X_i, \dots, X_{j-1}]$, with W-degree $d'_2 = d_2 - w_j\alpha$. If $d'_2 \geq d_1$, the result follows by induction. If $i = j$, then $m_2 = X_i^\alpha$, and $m_1 \coloneqq X_i^{d_1/w_i}$ has W-degree d_1 and divides m_2. So we can assume that $d'_2 < d_1$ and that $i < j$.

Since W is a reverse chain-divisible system of weights, w_{j-1} divides w_k for any k in $\{1, \dots, j-1\}$. Hence, since $m_2 \in \mathbb{K}[X_i, \dots, X_j]$ and $m'_2 \in \mathbb{K}[X_i, \dots, X_{j-1}]$, $d_2 \equiv 0\,[w_j]$ and $d'_2 \equiv 0\,[w_{j-1}]$. By hypothesis, d_1 is divisible by w_i, and in particular it is divisible by w_{j-1}. All in all, this shows that $d_1 - d'_2$ is divisible by w_{j-1}, and so it is divisible by w_j. Let

$$m_1 \coloneqq m'_2 \cdot X_j^{(d_1 - d'_2)/w_j}.$$

Then the monomial m_1 has W-degree d_1 and divides m_2.

(2 $\implies$ 1). Assume that W is a system of weights which is not reverse chain-divisible. We shall find integers $d_1 \leq d_2$ and a monomial m_2 with W-degree d_2 which is not divisible by any monomial of W-degree d_1.

Since W is not reverse chain-divisible, there exists i such that w_{i+1} does not divide w_i. In particular, $\gcd(w_i, w_{i+1}) < w_i$ and $\gcd(w_i, w_{i+1}) < w_{i+1}$. Without loss of generality, we may consider only the variables X_i, X_{i+1}. Let $d_1 = w_i w_{i+1}$, $d_2 = d_1 + \gcd(w_i, w_{i+1})$. By Paoli's lemma (see for example [Luc91, chap. 264] or the discussion after [NZM91, th. 5.1]), there exists exactly

$$\left\lfloor \frac{d_2}{w_i w_{i+1}} \right\rfloor = \left\lfloor 1 + \frac{\gcd(w_i, w_{i+1})}{w_i w_{i+1}} \right\rfloor = 1$$

couple of non-negative integers a, b such that $aw_i + bw_{i+1} = d_2$. Let m_2 be the monomial $X_i^a X_{i+1}^b$. The W-degree d_1 is divisible by w_i, and m_2 is not divisible by $X_1, \ldots, X_{i-1}$. The maximal divisors of m_2 are

$$\frac{m_2}{X_i} = X_i^{a-1} X_{i+1}^b \text{ with } W\text{-degree } d_2 - w_i = d_1 + \gcd(w_i, w_{i+1}) - w_i < d_1;$$
$$\frac{m_2}{X_{i+1}} = X_i^a X_{i+1}^{b-1} \text{ with } W\text{-degree } d_2 - w_{i+1} = d_1 + \gcd(w_i, w_{i+1}) - w_{i+1} < d_1.$$

As a consequence, m_2 is not divisible by any monomial of W-degree d_1. □

This proposition essentially states that the staircase of a W-homogeneous ideal is reasonably shaped when W is a reverse chain-divisible system of weights. For example, let W be a reverse chain-divisible system of weights, and let I be the ideal generated by all monomials of W-degree w_1 (that is, the least common multiple of the weights). Then the proposition proves that I contains all monomials of W-degree greater than w_1.

If on the other hand the system of weights is not reverse chain-divisible, this property needs not hold. For example, consider the algebra $\mathbb{K}[X_1, X_2, X_3]$ graded w.r.t. the system of weights $W = (3, 2, 1)$, the least common multiple of the weights being 6, and let I be the ideal generated by all monomials of W-degree 6. Consider the monomial $X_1 X_2^2$: it has W-degree 7, yet it is not divisible by any monomial with W-degree 6, and so it does not belong to the ideal I.

3.2.4. Genericity of regularity properties and W-compatibility

Proposition 3.17. *Let $m \leq n$ be two integers, $W = (w_1, \ldots, w_n)$ a system of weights, and $D = (d_1, \ldots, d_m)$ a system of W-degrees. Then*

- *the set of regular sequences,*
- *the set of sequences in Noether position with respect to the variables $X_1, \ldots, X_m$, and*
- *the set of sequences in simultaneous Noether position w.r.t. the order $X_1 > \cdots > X_m$*

are Zariski-open subsets of the affine space of W-homogeneous polynomials with W-degree D.

This states that the set of regular sequences, sequences in Noether position and sequences in simultaneous Noether position are Zariski-dense subsets if and only if they are not empty. Unfortunately, depending on the weights and the weighted degrees, there may exist no regular sequence, and thus no sequences in (simultaneous) Noether position either. For example, let $W = (2, 5)$ and $D = (4, 8)$, the only W-homogeneous sequence with W-degree D in $\mathbb{K}[X, Y]$ is (up to scalar multiplication) (X^2, X^4), and it is not regular. However, this is only the case for very specific systems of W-degrees for which there does not exist enough monomials to build non-trivial sequences.

Definition 3.18. Let $m \leq n$ be two integers, $W = (w_1, \dots, w_n)$ a system of weights, and $D = (d_1, \dots, d_m)$ a system of W-degrees. We say that D is *W-compatible* if there exists a regular W-homogeneous sequence in $\mathbb{K}[X_1, \dots, X_n]$ with W-degree D. We say that D is *strongly W-compatible* if for any $1 \leq i \leq m$, d_i is divisible by w_i.

Using these definitions, we can identify cases where the properties of being regular, in Noether position or in simultaneous Noether position are generic.

Proposition 3.19. *Let $m \leq n$ be two integers, $W = (w_1, \dots, w_n)$ a system of weights, and $D = (d_1, \dots, d_m)$ a system of W-degrees. For any $1 \leq i \leq m$, write $W_i \coloneqq (w_1, \dots, w_i)$ and $D_i \coloneqq (d_1, \dots, d_i)$. Write $A_{W,D}$ the affine space of W-homogeneous sequences of W-degree D. Then the following statements are true:*

1. *if D is W-compatible, then regular sequences form a Zariski-dense subset of $A_{W,D}$;*
2. *if D is W_m-compatible, then sequences in Noether position with respect to the variables $X_1, \dots, X_m$ form a Zariski-dense subset of $A_{W,D}$;*
3. *if D is strongly W-compatible, then D is W-compatible, W_m-compatible, and for any i, D_i is W_i-compatible;*
4. *if $m = n$, D is W-compatible and W is reverse chain-divisible, then, up to some reordering of the degrees, D is strongly W-compatible.*

Proof. The proofs of statements 1 and 2 follow the same technique: by Theorem 3.17, we know that the sets we consider are Zariski-open in $A_{W,D}$. So in order to prove the density, we only need to prove that they are non empty. Statement 1 is exactly the definition of the W-compatibility.

For statement 2, by W_m-compatibility, we know that there exists a W-homogeneous sequence $F = (f_1, \dots, f_m)$ with W-degree D in $\mathbb{K}[X_1, \dots, X_m]$, which is regular. As a consequence, the sequence $(f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ is regular, and from the characterization 4 of Noether position (prop. 1.44), this means that F is in Noether position with respect to the variables $X_1, \dots, X_m$.

In order to prove statement 3, let $1 \leq i \leq m$, we need to exhibit a regular sequences of length i in $\mathbb{K}[X_1, \dots, X_i]$. We may choose $F_i \coloneqq (X_1^{d_1/w_1}, \dots, X_i^{d_i/w_i})$, it is regular and each polynomial lies in $\mathbb{K}[X_1, \dots, X_i]$.

Finally, statement 4 is a consequence of Theorem 3.15. Let W be a reverse chain-divisible system of weights, and D a W-compatible system of W-degrees. Up to reordering, we can assume that the polynomials are ordered so that $d_1 \geq d_2 \geq \cdots \geq d_n$; this does not cancel the

W-compatibility. Let $F = (f_1, \dots, f_n)$ be a regular sequence, W-homogeneous with W-degree D. By Theorem 3.15, there exist polynomials $P_i(X_{i+1}, \dots, X_n)$ which are W-homogeneous with W-degree w_i, and such that F, under the change of variables $X_i = X'_i + P_i(X_{i+1}, \dots, X_n)$, is in simultaneous Noether position with respect to the order $X'_1 > X'_2 > \cdots > X'_n$.

From the characterization 4 of Noether position, this means in particular that for any $i \in \{1, \dots, n\}$, $f_i(X_1(X'_1, \dots, X'_i), \dots, X_n(X'_1, \dots, X'_i))$ belongs to a regular sequence, and thus is not zero. By definition of reverse chain-divisible weights, its W-degree d_i is a sum of multiples of w_i, and so it is itself a multiple of w_i. □

Remark 3.20. The statement 4 is a converse of 3 in the reverse chain-divisible case. In the non-reverse chain-divisible case, that converse is false: let $W = (3, 2)$, $D = (6, 5)$ and consider $F = (X^2 + Y^3, XY)$ in $\mathbb{K}[X, Y]$. The sequence F is in simultaneous Noether position w.r.t. the order $X > Y$, yet 5 is neither divisible by 3 nor by 2.

The weaker converse that if D is W-compatible, then D is W_m-compatible is also false: with the same weights and algebra, let $D = (5)$, the only polynomial with W-degree 5 is (up to scalar multiplication) $f = XY$. It is non-zero, so that (f) is a regular sequence, but (f, Y) is not regular, hence (f) is not in Noether position w.r.t X.

Remark 3.21. These examples lead to the following attempt at writing a general characterization of W-compatibility.

Let n be a positive integer, $W = (w_1, \dots, w_n)$ a system of weights, and $D = (d_1, \dots, d_n)$ a system of W-degrees. Further assume that

- for all $i \in \{1, \dots, n\}$, $\mathbb{K}[\mathbf{X}]_{d_i} \neq 0$
- the formal series

$$S_{D,W}(T) = \frac{\prod_{i=1}^{n}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})}$$

  is a polynomial.

Is D necessarily W-compatible?

The answer is *no*: take the system of weights $W = (3, 5, 11)$, and the system of W-degrees $D = (165, 19, 19)$. Note that 165 is the product of the weights, and 19 the sum of the weights. The series

$$S_{D,W}(T) = \frac{(1 - T^{165}) \cdot (1 - T^{19}) \cdot (1 - T^{19})}{(1 - T^3) \cdot (1 - T^5) \cdot (1 - T^{11})} = 1 + T^3 + \cdots + T^{184}$$

is a polynomial. But at W-degree 19, there are only 2 monomials, namely $X_1X_2X_3$ and $X_1^3X_2^2$, and they are not coprime, so we cannot form a regular sequence of W-degrees $(165, 19, 19)$.

Remark 3.22. Bézout's bound gives another necessary condition for D to be W-compatible in the zero-dimensional case. Indeed, if $n = m$ and D is W-compatible, let I be an ideal generated by a regular sequence of W-homogeneous polynomials with respective W-degree D. Then $\mathbb{K}[\mathbf{X}]/I$ is a finite-dimensional $\mathbb{K}$-vector space, and its dimension is given by

$$\dim_{\mathbb{K}}(\mathbb{K}[\mathbf{X}]/I) = \deg(I) = \frac{\prod_{i=1}^{n} d_i}{\prod_{i=1}^{n} w_i}.$$

Since this dimension is necessarily an integer, $\prod_{i=1}^{n} w_i$ divides $\prod_{i=1}^{n} d_i$.

In the homogeneous case, affine systems were studied by homogenization. In order to be able to study affine systems with a system of weights, we introduce the following definition:

Definition 3.23. Let $F = (f_1, \dots, f_m) \subset \mathbb{K}[X_1, \dots, X_n]$ be a sequence of polynomials, and let W be a system of weights. We say that F has a *W-homogeneous structure* if the highest W-degree components of $f_1, \dots, f_n$ form a regular sequence.

3.2.5. Characterization of the Hilbert series

Let $W = (w_1, \dots, w_n)$ be a reverse chain-divisible system of weights such that $w_n = 1$, and let $D = (d_1, \dots, d_n)$ be a system of W-degrees, such that for any $i \in \{1, \dots, n\}$, d_i is divisible by all of the w_j's. Let $R = \mathbb{K}[X_1, \dots, X_n]$ be a polynomial algebra graded with respect to W.

We use the notations below, following [Mor96]:

- $\delta_j = \sum_{i=1}^{j} (d_i - w_i)$;
- $\delta = \delta_n$, $\delta^* = \delta_{n-1}$;
- $\sigma = \min\left(\delta^*, \left\lfloor \frac{\delta}{2} \right\rfloor\right)$, $\sigma^* = \min\left(\delta_{n-2}, \left\lfloor \frac{\delta^*}{2} \right\rfloor\right)$;
- $\mu = \delta - 2\sigma$, $\mu^* = \delta^* - 2\sigma^*$.

Given a formal series $S(T) = \sum_{d=0}^{\infty} a_d T^d$, we also define

$$\begin{aligned} \Delta S(T) &= \sum_{d=0}^{\infty} (a_d - a_{d-1}) T^d \quad \text{(with the convention } a_{-1} = 0\text{)} \\ &= (1 - T) \cdot S(T) \end{aligned}$$

and

$$\int S = \sum_{d=0}^{\infty} (a_0 + \cdots + a_d) T^d = \frac{S(T)}{1 - T}.$$

Lemma 3.24. *Under the above notations and assumptions, the following properties hold.*

$$\begin{cases} \delta^* > \left\lfloor \frac{\delta}{2} \right\rfloor & \iff \quad d_n - \delta^* \leq 0 \\ \delta^* = \left\lfloor \frac{\delta}{2} \right\rfloor & \iff \quad 1 \leq \quad d_n - \delta^* \leq 2 \\ \delta^* < \left\lfloor \frac{\delta}{2} \right\rfloor & \iff \quad 3 \leq \quad d_n - \delta^* \end{cases} \tag{3.5}$$

$$\sigma = \left\lfloor \frac{\delta}{2} \right\rfloor \implies \mu = \delta \,[2] \in \{0, 1\} \tag{3.6}$$

$$0 \leq \mu < d_n \tag{3.7}$$

$$d_{n-1} \leq d_n \implies \sigma^* + \mu^* \leq \sigma \tag{3.8}$$

Proof. The proof of statements (3.5) and (3.7) can be found in [Mor96, Lemma 2.1]. This proof depends only on the value of w_n, and since we assume it to be 1, it is also valid in our setting. It also proves (3.6) as a side-result.

For the statement (3.8), we proceed by case disjunction on the values of σ.

- If $\sigma = \delta^*$:

$$\sigma^* + \mu^* = \delta^* - \sigma^* \leq \delta^* = \sigma.$$

- If $\sigma = \lfloor \delta/2 \rfloor$, then $\sigma = \lfloor (\delta^* + d_n - 1)/2 \rfloor$ which implies $2\sigma = \delta^* + d_n - 1 - \mu$ and $\mu = \delta\,[2] \in \{0, 1\}$ (from statement (3.6)). Now consider the possible values of σ^*:
  - if $\sigma^* = \lfloor \delta^*/2 \rfloor$, then $\mu^* = \delta^*\,[2]$, and thus $2\sigma = 2\sigma^* + \mu^* + d_n - 1 - \mu$. It implies that $d_n - 1 - \mu + \mu^*$ is even. We shall prove that it is greater than or equal to 0.

    From statement (3.7), $d_n - 1 - \mu \geq 0$, so if $\mu^* = 0$, we are done. If $\mu^* = 1$, by parity $d_n - 1 - \mu$ is odd, and thus $d_n - 1 - \mu \geq 1 = \mu^*$.

    It implies that:

$$2\sigma = 2\sigma^* + \mu^* + d_n - 1 - \mu \geq 2\sigma^* + 2\mu^*;$$

  - otherwise, $\sigma^* = \delta^{**}$, and in that case

$$\sigma^* + \mu^* = \delta^* - \sigma^* = \delta^* - \delta^{**} = d_{n-1} - w_{n-1}$$

    which implies that:

$$d_n - 1 \geq \sigma^* + \mu^* \text{ (since } w_{n-1} \geq w_n \text{ and } d_{n-1} \leq d_n)$$

    and

$$\delta^* = \delta^{**} + d_{n-1} - w_{n-1} \geq \sigma^* + \mu^*.$$

    So we have:

$$\begin{aligned} 2\sigma &= \delta^* + d_n - 1 - \mu \\ &\geq \sigma^* + \mu^* + \sigma^* + \mu^* - \mu. \end{aligned}$$

    Recall that $\mu \in \{0, 1\}$, so by parity, $2\sigma \geq 2\sigma^* + 2\mu^*$, hence $\sigma \geq \sigma^* + \mu^*$. □

The following theorem describes the shape of the Hilbert series of a zero-dimensional complete intersection. It states that it is a self-reciprocal (or palindromic) polynomial, that is a polynomial with symmetrical coefficients, and that these coefficients increase at small degrees, then station, then decrease again. Furthermore, between every strict increase, they reach a step, which has width w_{n-1}. For an example, see figure 3.3, where the width of the steps is 3, and the width of the central plateau is 5.

This is a generalization of a known result in the homogeneous case (Proposition 1.40) which has been proved for example in [Mor96, prop. 2.2] (we will follow that proof for the weighted case). The result is simpler in the homogeneous case: there is no such step in the growth of the coefficients, and they are strictly increasing, then stationary, then strictly decreasing.

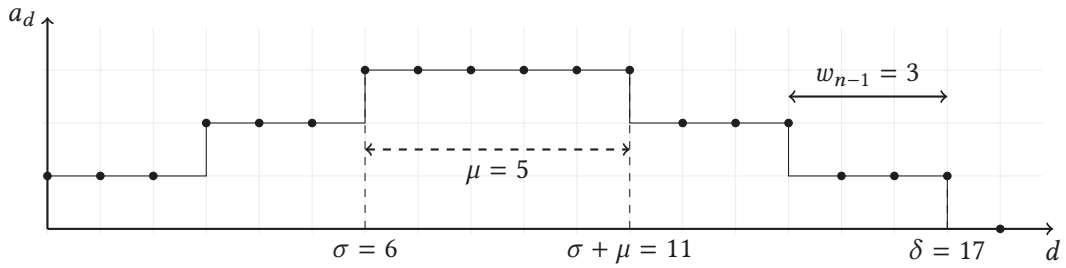

Figure 3.3: Shape of the Hilbert series of a W-homogeneous complete intersection for $W = (3, 3, 1)$ and $D = (9, 6, 3)$

Theorem 3.25. *Let $W = (w_1, \dots, w_n)$ be a reverse chain-divisible system of weights, and $D = (d_1, \dots, d_n)$ a system of degrees such that for any $i \in \{1, \dots, n\}$, d_i is divisible by w_1. Consider the formal series*

$$S_{W,D}(T) = \frac{\prod_{i=1}^{n}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})} = \sum_{d=0}^{\delta} a_d T^d$$

The series $S_{W,D}$ is a self-reciprocal polynomial in T (i.e. for any $d \leq \delta$, $a_d = a_{\delta - d}$) and its coefficients satisfy the inequalities:

$$\begin{aligned} \forall d \in \{0, \dots, \sigma - 1\}, &\quad a_d \leq a_{d+1} \\ \forall d \in \{\sigma, \dots, \sigma + \mu - 1\}, &\quad a_d = a_{d+1} \\ \forall d \in \{\sigma + \mu, \dots, \delta\}, &\quad a_d \geq a_{d+1} \end{aligned}$$

Furthermore, if $d < \sigma$ (resp. $d > \sigma + \mu$), the coefficients increase (resp. decrease) with steps, and these steps have width w_{n-1}:

$$\forall d \in \{0, \dots, \sigma - 1\},\ a_d - a_{d-1} \begin{cases} > 0 & \text{if } w_{n-1} \text{ divides } d \\ = 0 & \text{otherwise.} \end{cases}$$

Proof. We adapt the proof from [Mor96, Prop. 2.2] for the homogeneous case to the weighted case. Up to permutation of the d_i's, we can assume that for any i, $d_i \geq d_{i-1}$. We proceed by induction on n. The result for the case $n = 1$ is a consequence of the homogeneous case, since $w_n = 1$.

Let $n > 1$. Let $\overline{W}^* = (w_1/w_{n-1}, \dots, w_{n-1}/w_{n-1})$ and $\overline{D}^* = (d_1/w_{n-1}, \dots, d_{n-1}/w_{n-1})$, and consider the series

$$\overline{S}^* := S_{W^*,D^*} = \frac{\prod_{i=1}^{n-1}(1 - T^{d_i/w_{n-1}})}{\prod_{i=1}^{n-1}(1 - T^{w_i/w_{n-1}})} = \sum_{d=0}^{\delta} \overline{a}_d^* T^d.$$

Write $\overline{S}^*$ the Hilbert series of $\overline{F}^*$, with generic coefficient $\overline{a}_d^*$. The system $\overline{F}^*$ satisfies the hypotheses of the theorem (from the caracterization 3 of the Noether position 1.44), so by induction hypothesis, $\overline{S}^*$ has the shape predicted by the theorem.

The Hilbert series S can be computed from $\overline{S}^*$ with

$$S(T) = \frac{1 - T^{d_n}}{1 - T}\overline{S}^*(T^{w_{n-1}}) = (1 - T^{d_n}) \cdot \int \overline{S}^*(T^{w_{n-1}}),$$

and so for any d, we have:

$$\begin{aligned} a_d &= a^*_{d-d_n+1} + \cdots + a^*_d \\ a'_d &:= a_d - a_{d-1} = a^*_d - a^*_{d-d_n} \end{aligned}$$

where

$$a^*_d = \begin{cases} \overline{a}^*_{\overline{d}} & \text{if } d = \overline{d}w_{n-1} \\ 0 & \text{otherwise.} \end{cases}$$

This proves that the polynomial is self-reciprocal. Indeed:

$$\begin{aligned} a_{\delta-d} &= a^*_{\delta-d-d_n+1} + \cdots + a^*_{\delta-d} \\ &= a^*_{d-d_n+1} + \cdots + a^*_d \text{ since, by induction hypothesis, } \overline{S}^* \text{ is self-reciprocal} \\ &= a_d \end{aligned}$$

To prove the properties regarding the sign of $a'_d = a_d - a_{d-1}$, we shall consider two cases, according to the value of d_n.

- If $d_n \geq \delta^* + 1$, then from statement (3.5) in Lemma 3.24, and the definition of σ and μ, $\sigma = \delta^*$ and $\sigma + \mu = d_n - 1$. Let $0 \leq d \leq \sigma$, then $d \leq \delta^* < d_n$, so:

  $$a'_d = a^*_d = \begin{cases} \overline{a}^*_{d/w_{n-1}} > 0 & \text{if } w_{n-1} \text{ divides } d; \\ 0 & \text{otherwise.} \end{cases}$$

  Let $d \in \{\sigma + 1, \ldots, \sigma + \mu\}$, that implies that $\delta^* < d \leq d_n - 1$, so:

  $$a'_d = a^*_d = 0 \text{ (since } \delta^* \text{ is the degree of } S^*).$$

- If $d_n \leq \delta^*$, then from statement (3.5) again, $\sigma = \lfloor \delta/2 \rfloor$ and $\mu = \delta\,[2]$. Let $d \leq \sigma$. We want to prove that $a_d - a_{d-1}$ is greater or equal to zero, depending on whether d is divisible by w_{n-1}. We shall consider two ranges of values for d:
  - if $d \leq \sigma^* + \mu^*$, then $d - d_n \leq \sigma^* + \mu^* - d_n < \sigma^*$ (since $\mu^* < d_n$). Recall that $a'_d = a^*_d - a^*_{d-d_n}$. By hypothesis, d_n is divisible by w_{n-1}, and so, either both d and $d - d_n$ are divisible by w_{n-1}, or both are not. Thus,

    $$a'_d \begin{cases} > 0 & \text{if both } d \text{ and } d - d_n \text{ are divisible by } w_{n-1} \\ = 0 & \text{if neither } d \text{ nor } d - d_n \text{ is divisible by } w_{n-1}; \end{cases}$$

– if $\sigma^*+\mu^* < d \leq \sigma$, then $2d \leq 2\sigma \leq \delta$; by definition, $\delta = \delta^*+d_n-1$, so $d-d_n < \delta^*-d$; furthermore, $\delta^* - d < \delta^* - (\sigma^* + \mu^*) = \sigma^*$, so in the end:

$$d - d_n < \delta^* - d < \sigma^*.$$

Since, by construction, δ^* is divisible by w_{n-1}, the same reasoning as before yields that

$$a'_d = a^*_d - a^*_{d-d_n} = a^*_{\delta^*-d} - a^*_{d-d_n}$$

and

$$a'_d \begin{cases} > 0 & \text{if both } \delta^* - d \text{ and } d - d_n \text{ are divisible by } w_{n-1}; \\ = 0 & \text{if neither } \delta^* - d \text{ nor } d - d_n \text{ is divisible by } w_{n-1}. \end{cases}$$

Still assuming that $d_n \leq \delta^*$, let now $d \in \{\sigma + 1, \ldots, \sigma + \mu\}$. We want to prove that $a_d - a_{d-1} = 0$. If $\mu = 0$ there is nothing to prove, so assume that $\mu = 1$ and $d = \sigma + 1$. But then $\sigma + 1 - d_n = \delta - \sigma - d_n = \delta^* - \sigma$, and so by symmetry, $a'_d = a^*_{\sigma+1} - a^*_{\sigma+1-d_n} = 0$. □

Remark 3.26. The hypothesis that the weights are reverse chain-divisible is necessary. As a counter-example, let $W = (3, 2, 2)$ and $D = (6, 6, 6)$. Then the Hilbert series of a complete intersection of W-degree D is illustrated in Figure 3.4. It is self-reciprocal, but the coefficients do not vary as predicted by Theorem 3.25.

The hypothesis that each of the W-degrees should be divisible by w_1 is also necessary. As a counter-example, let $W = (4, 2, 1)$ and $D = (8, 8, 2)$. Then the Hilbert series of a complete intersection of W-degree D is illustrated in Figure 3.5: the width of the steps is greater than w_{n-1}. Furthermore, following the proof, the parameters for this series should be defined by $\sigma = \lfloor \delta/2 \rfloor$ and $\mu = \delta\,[2]$, where $\delta = 11$, so that $\sigma = 5$ and $\mu = 1$. However, we cannot reorder the degrees such that $d_3 \geq d_2 \geq d_1$, and we cannot deduce from statement (3.8) in Lemma 3.24 that $\sigma^* + \mu^* \leq \sigma$: indeed, we have $\sigma = 4$ but $\sigma^* + \mu^* = 6$.

However, the fact that the Hilbert series is self-reciprocal for complete intersections is true even for general system of weights, and is a consequence of the Gorenstein property of complete intersections (see [Eis95, Chap. 21]; this property is also central to the proof of Theorem 3.42).

3.3. W-Degree of regularity of regular sequences

3.3.1. Macaulay's bound in dimension zero

Let $W = (w_1, \ldots, w_n)$ be a system of weights, and $F = (f_1, \ldots, f_m)$ a regular sequence of W-homogeneous polynomials, with respective W-degrees $d_1, \ldots, d_m$. Further assume that the set of solutions is zero-dimensional, that is $m = n$. We denote by I the weighted homogeneous ideal generated by F.

One can read on the Hilbert series a weighted version of Macaulay's bound:

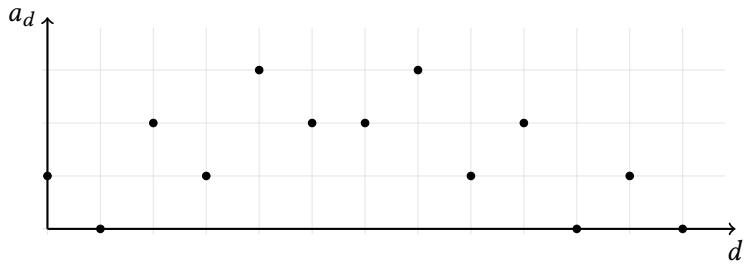

Figure 3.4: Hilbert series of a weighted homogeneous complete intersection with $W = (3, 2, 2)$ and $D = (6, 6, 6)$ — The weights are not reverse chain-divisible.

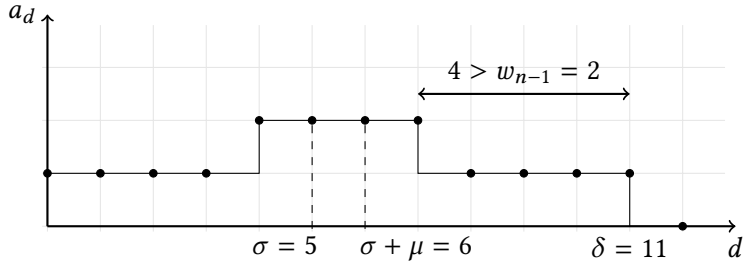

Figure 3.5: Hilbert series of a weighted homogeneous complete intersection with $W = (4, 2, 1)$ and $D = (8, 8, 2)$ — The W-degrees are not all divisible by w_1.

Theorem 3.27 (Macaulay's bound, weighted case)**.** *With the same notations and hypotheses, the W-degree of regularity of F is bounded by*

$$d_{\mathrm{reg},W}(F) \leq i_{reg}(I) + \max\{w_i\} = \sum_{i=1}^{n} d_i - w_i + \max\{w_i\}. \tag{3.9}$$

Proof. As in the homogeneous case, since I is zero-dimensional, the Hilbert series is a polynomial with degree i_{reg}. This means that all monomials of W-degree greater than i_{reg} are in the ideal, and as such, that the leading terms of the W-GRevLex Gröbner basis of F need to divide all the monomials of W-degree greater than i_{reg}.

Let $f \in \mathbb{K}[\mathbf{X}]$, W-homogeneous with W-degree at least $i_{\mathrm{reg}} + \max\{w_i\} + 1$. If it belongs to a reduced Gröbner basis of I, then none of the proper divisors of his leading term are in $\mathsf{LT}(I)$, but they necessarily have W-degree at least $i_{\mathrm{reg}} + 1$, so we reached a contradiction. Hence all polynomials in a reduced W-GRevLex Gröbner basis of I have W-degree at most $i_{\mathrm{reg}} + \max\{w_i\}$. □

In the homogeneous case, this bound is sharp, and is actually reached for generic system, but it is not true in the weighted case. In particular, it appears that the W-degree of regularity depends on the order set on the variables (see Table 3.1 for some examples).

3.3.2. Sharper bound on the degree of regularity

The following theorem is an improvement over the previous bound, under the additional assumption that the system is in simultaneous Noether position. Recall that this property is generic, and that for reverse chain-divisible systems of weights, it is always true for regular sequences, up to a weighted homogeneous change of coordinates.

Theorem 3.28. *Let $W = (w_1, \dots, w_n)$ be a (not necessarily reverse chain-divisible) system of weights and $D = (d_1, \dots, d_n)$ be a strongly W-compatible system of W-degrees. Further assume that for any $j \in \{2, \dots, n\}$, $d_j \geq w_{j-1}$. Let $F = (f_1, \dots, f_n)$ be a system of W-homogeneous polynomials, with W-degree D, and assume that F is in simultaneous Noether position for the variable ordering $X_1 > X_2 > \cdots > X_n$. Then the W-degree of regularity of F is bounded by*

$$d_{\mathrm{reg},W}(F) \leq \sum_{i=1}^{n}(d_i - w_i) + w_n. \tag{3.10}$$

Proof. We prove this by induction on n. If $n = 1$, we simply have one W-homogeneous polynomial to consider, and so $d_{\mathrm{reg},W} = d_1$.

Assume now that $n > 1$. We consider the system F^* defined by:

$$F^* = \big(f_1(X_1, \dots, X_{n-1}, 0), \dots, f_{n-1}(X_{n-1}, \dots, X_{n-1}, 0)\big).$$

This system is W^*-homogeneous, for $W^* := (w_1, \dots, w_{n-1})$. From the characterization 3 of Noether position, the sequence F^* is in simultaneous Noether position. As a consequence, the induction hypothesis applies to F^*, and the W^*-degree of regularity of F^* is bounded by

$$d_{\mathrm{reg},W^*}(F^*) \leq \sum_{i=1}^{n-1}(d_i - w_i) + w_{n-1}.$$

Let δ be the degree of the Hilbert series of F, that is $\delta = \sum_{i=1}^{n}(d_i - w_i)$. We want to prove that $d_{\mathrm{reg}} \leq \delta + w_n$, i.e. that the Gröbner basis of F need not contain any polynomial with W-degree greater than $\delta + w_n$. Equivalently, let μ be a monomial with W-degree $d > \delta + w_n$. We will prove that μ is strictly divisible by a monomial in the initial ideal generated by F.

Write $\mu = X_n^{\alpha} \cdot \mu'$, with $\mu' \in \mathbb{K}[X_1, \dots, X_{n-1}]$, and proceed by induction on α:

- if $\alpha = 0$, then $\mu \in \mathbb{K}[X_1, \dots, X_{n-1}]$. By assumption $d_n \geq w_{n-1}$, hence:

$$\delta + w_n = \delta^* + w_{n-1} - w_{n-1} + d_n - w_n + w_n \geq d^*_{\mathrm{reg}} + d_n - w_{n-1} \geq d^*_{\mathrm{reg}}; \tag{3.11}$$

  and so μ has W-degree greater than d^*_{reg}. By induction hypothesis, μ is strictly divisible by a monomial in the initial ideal generated by F^*.

- If $\alpha > 0$, then consider $\mu'' = X_n^{\alpha-1}\mu'$: it is a strict divisor of μ. Furthermore, since $\deg(\mu) > \delta + w_n$, then $\deg(\mu'') = \deg(\mu) - w_n > \delta$. Recall that δ is by definition the degree of the Hilbert series of the ideal generated by F, so μ lies in that ideal. □

Remark 3.29. The hypothesis stating that for any i, $d_i \geq w_{i-1}$ is necessary. For example, let $W = (2,1)$, $D = (2,1)$ and the system $F = (X, Y)$ in $\mathbb{K}[X, Y]$, it is W-homogeneous with W-degree D and in simultaneous Noether position. This system has Hilbert series 1 (the quotient vector span is generated by $\{1\}$), which has degree $\delta = 0$. But the Gröbner basis of the system is given by F itself, and contains X, with W-degree 2.

More generally, without that hypothesis, we obtain the following bound for $d_{\mathrm{reg},W}(F)$:

$$d_{\mathrm{reg},W}(F) \leq \max\left\{\sum_{i=1}^{k}(d_i - w_i) + w_k : k \in \{1,\ldots,n\}\right\},$$

and the proof is the same as that of Theorem 3.28, with the weaker induction hypothesis that $d_{\mathrm{reg},W}(F) \leq \max\left(\delta + w_n, d_{\mathrm{reg},W^*}(F^*)\right)$, which does not need inequality (3.11).

Remark 3.30. We give examples of the behavior of both bounds in Table 3.1: we give the degree of regularity of a generic W-homogeneous system of W-degree D, and show how this degree of regularity varies if we change the order of the weights W.

Remark 3.31. Theorem 3.28 gives an indication as to how to choose the order of the variables. Generically, in order to compute a W-GRevLex Gröbner basis of the system, the complexity estimates will be better if we set the variables in decreasing weight order.

3.3.3. Conjectured exact formula

While the new bound (3.10) is not sharp in full generality, it is sharp whenever $w_n = 1$. We conjecture that the sharp formula is the following.

Conjecture 3.32. Let $W = (w_1,\ldots,w_n)$ be a system of weights, and $D = (d_1,\ldots,d_n)$ a strongly W-compatible system of W-degrees. Let $F \in \mathbb{K}[X_1,\ldots,X_n]$ be a generic system of W-homogeneous polynomials. Let $\delta = \sum_{i=1}^{n}(d_i - w_i)$ be the degree of the Hilbert series of $\langle F\rangle$, and let d_0 be defined as

$$d_0 = \begin{cases} \delta + 1 & \text{if there exists } i \text{ such that } w_i = 1 \\ \delta - g & \text{otherwise}, \end{cases}$$

where g is the Frobenius number of W (that is, the greatest W-degree at which the set of monomials is empty). In other words, d_0 is the degree of the first "unexpected" zero coefficient in the Hilbert series (by definition of the degree in the first case, and by self-reciprocality of the Hilbert series in the second case).

Then the degree of regularity of F is the first multiple of w_n greater than d_0:

$$d_{\mathrm{reg}} = w_n \left\lceil \frac{d_0}{w_n} \right\rceil. \tag{3.12}$$

Table 3.1: Macaulay's bound on the degree of regularity of generic weighted homogeneous systems

W	D	d_{reg}	Bound (3.9)	Bound (3.10)
$(3,2,1)$	$(6,6,6)$	13	15	13
$(3,1,2)$	$(6,6,6)$	14	15	14
$(1,2,3)$	$(6,6,6)$	15	15	15

3.3.4. Positive-dimensional regular sequences

In the positive-dimensional case, the main complexity parameter that we study is the degree of regularity of the system. Under some Noether position assumptions, the bounds that we obtained for zero-dimensional systems apply.

As before, given $F = (f_1, \dots, f_m)$ we define $F_{\text{ext}} = (f_1, \dots, f_m, X_{m+1}, \dots, X_n)$ and $\theta_m(F) = F(X_1, \dots, X_m, 0, \dots, 0)$.

Theorem 3.33. *Let $F = (f_1, \dots, f_m)$ be a W-homogeneous polynomial system with respective W-degree $(d_1, \dots, d_m)$. If F is in Noether position with respect to $X_{m+1}, \dots, X_n$, then*

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - w_i) + \max_{1 \leq j \leq m} \{w_j\}.$$

If F is in simultaneous Noether position with respect to the order $X_1 > \cdots > X_m$, then

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - w_i) + w_m.$$

Proof. The proof is analogous to that of Theorem 2.47. □

3.4. Overdetermined systems

As in the homogeneous case, we want to be able to study systems with more equations than unknowns. For this purpose, we want to generalize the notion of semi-regularity to the weighted case.

We will use the same algebraic definition as in the homogeneous case (Definition 1.45). However, the characterization of Proposition 1.46 does not hold in full generality.

We prove that it does hold in the case of reverse chain-divisible systems of weights however. We then give some consequences on the degree of regularity of the ideal, and show that Fröberg's conjecture is true if $m = n + 1$, as in the homogeneous case.

3.4.1. Definition of semi-regularity

Let n and m be two integers, $m \geq n$, $W = (w_1, \dots, w_n)$ a system of weights, and $D = (d_1, \dots, d_m)$ a system of W-degrees. Let $F = (f_1, \dots, f_m)$ be a system of W-homogeneous polynomials with W-degree D. For any $i \in \{i, \dots, n\}$, write $F_i = (f_1, \dots, f_i)$.

Definition 3.34 (Semi-regularity)**.** We say that F is *semi-regular* if, for any $i \in \{1, \dots, m\}$ and for any $d \in \mathbb{N}$, the linear map given by the multiplication by f_i:

$$s_{i,d} : \left(\mathbb{K}[X_1, \dots, X_n]/\langle F_{i-1}\rangle\right)_d \xrightarrow{\cdot f_i} \left(\mathbb{K}[X_1, \dots, X_n]/\langle F_{i-1}\rangle\right)_{d+d_i}$$

is full-rank (either injective or surjective).

Furthermore, let

$$S_{D,W}(T) = \frac{\prod_{i=1}^{m}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})} = \sum_{d=0}^{\infty} a_d T^d.$$

We say that F *has a semi-regular Hilbert series* if the Hilbert series of F is equal to $\lfloor S_{D,W}(T) \rfloor$, that is the series truncated at the first coefficient less than or equal to zero.

The motivation behind these definitions is that in the homogeneous case, they are equivalent (Proposition 1.46):

Proposition 3.35. *If $W = (1, \dots, 1)$, the following conditions are equivalent:*

1. *the system F is semi-regular;*
2. *for any $1 \leq i \leq n$, the system F_i has a semi-regular Hilbert series.*

For weighted homogeneous systems, the converse implication ($2 \implies 1$) is still true:

Proposition 3.36. *Let F be a W-homogeneous system such that, for any $1 \leq i \leq n$, the system F_i has a semi-regular Hilbert series. Then F is semi-regular.*

Proof. We prove this by induction on the number m of polynomials. The initial case is $m = n$, and it is a direct consequence of the characterization of a regular sequence.

Assume $m > n$. Write $R^* = \mathbb{K}[X_1, \dots, X_n]/\langle f_1, \dots, f_{m-1}\rangle$, and for any $d \in \mathbb{N}$, consider the multiplication map

$$s_{m,d} = R^*_d \xrightarrow{\cdot f_m} R^*_{d+d_m}$$

Let $K_{m,d} = \ker(s_{m,d})$. Write $S(T)$ the Hilbert series of F, a_d its coefficient at degree d, δ its degree, $H(T) = \left(\prod_{i=1}^{m}(1 - T^{d_i})\right) / \left(\prod_{i=1}^{n}(1 - T^{w_i})\right)$, b_d its coefficient at degree d, and $S^*(T)$, a^*_d, δ^*, $H^*(T)$ and b^*_d their counterparts with $m - 1$ polynomials. From the exact sequence

$$0 \longrightarrow K_{m,d} \longrightarrow R^*_d \xrightarrow{s_{m,d}} R^*_{d+d_m} \longrightarrow R_{d+d_m} \longrightarrow 0,$$

we know that the following identity holds

$$a_{d+d_m} = a^*_{d+d_m} - a^*_d + \dim(K_{m,d}).$$

We want to prove that either $a_{d+d_m} = 0$ or $\dim(K_{m,d}) = 0$. Assume that $a_{d+d_m} > 0$. This means that $d + d_m \leq \delta$ and $a_{d+d_m} = b_{d+d_m}$, so:

$$\begin{aligned} a_{d+d_m} &= a_{d+d_m} - a^*_d + \dim(K_{m,d}) \\ &= b_{d+d_m} \\ &= b^*_{d+d_m} - b^*_d \text{ by definition of } H(T) \\ &= a^*_{d+d_m} - a^*_d \text{ since } \delta^* \geq \delta. \end{aligned}$$

Thus $K_{m,d} = 0$. □

3.4.2. Characterization with the Hilbert series

In this section, we prove that for reverse chain-divisible systems of weights, semi-regular sequences have a semi-regular Hilbert series. First, we characterize the shape of semi-regular Hilbert series, by extending Theorem 3.25 to the overdetermined case.

Theorem 3.37. *Let $m \geq n \geq 0$ be two integers. Let $W = (w_1, \dots, w_n)$ be a reverse chain-divisible system of weights, and let $D = (d_1, \dots, d_m)$ be a system of W-degrees such that $d_1, \dots, d_m$ are all divisible by w_1. Write*

$$S_{D,W}(T) = \frac{\prod_{i=1}^{m}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})} = \sum_{d=0}^{\infty} a_d T^d.$$

Then there exist W-degrees σ, δ such that

$$\forall\, d \in \{1, \dots, \sigma\},\ a_d \geq a_{d-1} \tag{σ1}$$
$$a_\sigma > a_{\sigma-1} \tag{σ2}$$
$$\forall\, d \in \{\sigma + 1, \dots, \delta\},\ a_d \leq a_{d-1} \tag{σ3}$$
$$a_\delta > 0,\ a_{\delta+1} \leq 0. \tag{δ1}$$

Furthermore, if $m > n$, let $D^* = (d_1, \dots, d_{m-1})$ and define δ^* as above for the series $S_{D^*,W}$. Then the following statements hold:

$$\begin{cases} \forall\, d \in \{\delta + 1, \dots, \delta^*\},\ a_d \leq 0 & \text{if } n = 0 \\ \forall\, d \in \{\delta + 1, \dots, \delta^* + d_m\},\ a_d \leq 0 & \text{if } n > 0. \end{cases} \tag{δ2}$$

If $n > 0$, let $W^* = (w_1, \dots, w_{n-1})$, and let δ' be the degree of $\lfloor S_{D,W^*}(T) \rfloor$. If $n = 0$, let $\delta' = 0$. Then the following equality holds

$$\sigma = \delta'. \tag{σ4}$$

Proof. We prove the theorem by induction on n, and for any given n, by induction over m. The base cases are:

- $n = 0$, $m \geq 0$: then $S_{D,W}(T) = 1 - a_k T^k + O(T^{k+1})$ with $a_k > 0$, and we can conclude, taking $\delta = 0$ and $\sigma = 0$.
- $n = m > 0$: then this is a consequence of Theorem 3.25 (shape of the Hilbert series of a complete intersection).

Assume that $m > n > 0$. Let $D^* = (d_1, \dots, d_{m-1})$, $W^* = (w_1, \dots, w_{n-1})$, and write:

$$S(T) \coloneqq S_{D,W}(T) = \sum_{d=0}^{\infty} a_d T^d;$$
$$S^*(T) \coloneqq S_{D^*,W}(T) = \sum_{d=0}^{\infty} a^*_d T^d.$$

The derivatives of these series are

$$\Delta S(T) = S_{D,W^*}(T) = \sum_{d=0}^{\infty} a'_d T^d;$$
$$\Delta S^*(T) = S_{D^*,W^*}(T) = \sum_{d=0}^{\infty} a'^*_d T^d.$$

Furthermore, let $w = w_{n-1}$, $\overline{W^*} = (w_1/w, \dots, w_{n-1}/w)$ and $\overline{D^*} = (d_1/w, \dots, d_{n-1}/w)$, and consider the series

$$\overline{\Delta S}(T) = S_{\overline{D},\overline{W^*}}(T) = \sum_{d=0}^{\infty} \overline{a'} dT^d;$$
$$\overline{\Delta S^*}(T) = S_{\overline{D^*},\overline{W^*}}(T) \sum_{d=0}^{\infty} \overline{a'^*} dT^d.$$

In particular,

$$\Delta S(T) = \overline{\Delta S}(T^w) \text{ and } \Delta S^*(T) = \overline{\Delta S^*}(T^w).$$

All the series S^*, $\overline{\Delta S}$ and $\overline{\Delta S^*}$ satisfy the induction hypothesis. The W-degrees for which the coefficients of S^* satisfy properties ($\sigma 1$)-($\sigma 4$) and ($\delta 1$)-($\delta 2$) are denoted by σ^* and δ^*. We write $\overline{\sigma'}$, $\overline{\delta'}$, $\overline{\sigma'^*}$, $\overline{\delta'^*}$ the respective values of the W-degrees for which these properties apply to $\overline{\Delta S}$ and $\overline{\Delta S^*}$.

From $S(T) = (1 - T^{d_m})S^*(T)$, we deduce the recurrence relation

$$a_d = a^*_d - a^*_{d-d_m}.$$

Since S^* satisfies the induction hypothesis, we know that there exists a degree δ such that

$$\begin{cases} \forall\, d \in \{0, \dots, \delta\} & a^*_d > a^*_{d-d_m} \\ \forall\, d \in \{\delta + 1, \dots, \delta^* + d_m\} & a^*_d \le a^*_{d-d_m}. \end{cases}$$

This proves statements ($\delta 1$) and ($\delta 2$). As a side result, since $a^*_\delta > a^*_{\delta - d_m}$, we also deduce that

$$\delta - d_m < \sigma^*. \tag{3.13}$$

Let $\sigma = \delta'$, we prove that it satisfies equations ($\sigma 1$), ($\sigma 2$) and ($\sigma 3$). We need to evaluate the sign of $a_d - a_{d-1}$, depending on d. The generating series of $a_d - a_{d-1}$ is:

$$(1 - T)S(T) = (1 - T) \cdot \frac{\prod_{i=1}^{m}(1 - T^{d_i})}{\prod_{i=1}^{n}(1 - T^{w_i})} = \Delta S(T) \text{ since } w_n = 1.$$

In other words, $a_d \ge a_{d-1}$ if and only if $a'_d \ge 0$, which proves statements ($\sigma 1$) and ($\sigma 2$), by definition of δ':

$$\forall\, d \in \{0, \dots, \sigma\},\ a_d - a_{d-1} = a'_d \ge 0$$
$$a_\sigma - a_{\sigma-1} = a'_\sigma = a'_{\delta'} > 0.$$

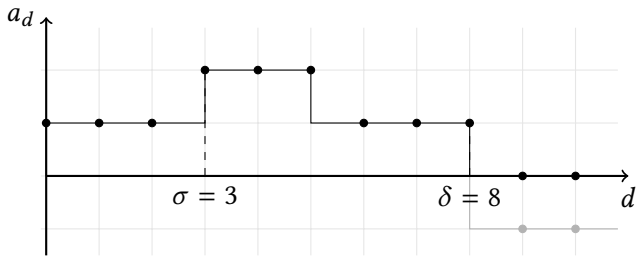

Figure 3.6: Shape of the Hilbert series of a semi-regular W-homogeneous sequence with $W = (3, 3, 1)$ and $D = (12, 9, 6, 6, 3)$

To finish the proof, we need to prove that for any $d \in \{\delta' + 1, \dots, \delta\}$, $a'_d \leq 0$.

From the induction hypothesis (statement ($\sigma 4$)) applied to S^*, we know that $\delta'^* = \sigma^*$. Moreover, $\overline{\Delta S}$ satisfies the induction hypothesis, and statement ($\delta 2$) yields that:

$$\forall\, \overline{d} \in \{\overline{\delta'} + 1, \dots, \overline{\delta'^*}\},\ \overline{a'_d} \leq 0.$$

As a consequence, since $\sigma^* = \delta'^* = w\overline{\delta'^*}$:

$$\forall\, d \in \{\delta' + 1, \dots, \sigma^*\},\ a'_d = \begin{cases} \overline{a'_{\overline{d}}} \leq 0 & \text{if } d = w\overline{d}; \\ 0 & \text{otherwise.} \end{cases} \tag{i}$$

Now assume that $\sigma^* < d \leq \delta$. We can write a'_d as

$$\begin{aligned} a'_d = a_d - a_{d-1} &= a^*_d - a^*_{d-d_m} - a^*_{d-1} + a^*_{d-d_m-1} \\ &= (a^*_d - a^*_{d-1}) - (a^*_{d-d_m} - a^*_{d-d_m-1}) = a'^*_d - a'^*_{d-d_m}. \end{aligned}$$

Since $a_d \leq a^*_d$ for any d, we necessarily have $\delta \leq \delta^*$, hence $\sigma^* < d \leq \delta^*$. Then, by induction hypothesis (statement ($\sigma 3$)), we know that $a^*_d - a^*_{d-1} \leq 0$. Additionally, equation (3.13) and induction hypothesis (statement ($\sigma 1$)) together yield that $a^*_{d-d_m} - a^*_{d-d_m-1} \geq 0$, so we conclude that

$$\forall\, d \in \{\sigma^* + 1, \dots, \delta\},\ a'_d \leq 0. \tag{ii}$$

Sticking the ranges of statements (i) and (ii) together, we prove statement ($\sigma 3$) which completes the proof. □

Using this description of semi-regular Hilbert series, we now prove that for reverse chain-divisible systems of weights, semi-regular sequences have a semi-regular Hilbert series. As an illustration, Figure 3.6 shows the coefficient of a semi-regular Hilbert series. The black dots correspond to the actual coefficients, and the gray dots are the coefficients which were truncated away.

Theorem 3.38. *Let $m \geq n \geq 0$ be two integers, $W = (w_1, \dots, w_n)$ be a reverse chain-divisible system of weights and $D = (d_1, \dots, d_m)$ be a system of W-degrees such that $d_1, \dots, d_n$ are all*

divisible by w_1. Let $F = (f_1, \ldots, f_m)$ be a system of W-homogeneous polynomials, with respective W-degree D.

If F is a semi-regular sequence, then F has a semi-regular Hilbert series.

Proof. We proceed by induction on m. If $m = n$, then the result is a consequence of the characterization of regular sequences.

Assume that $m > n$. We consider the series $S(T) = S_{D,W}(T)$ with generic coefficient a_d, $S^*(T) = S_{D^*,W}(T)$ with generic coefficient a^*_d, $H(T)$ the Hilbert series of F with generic coefficient b_d, and $H^*(T)$ the Hilbert series of $F^* := (f_1, \ldots, f_{m-1})$ with generic coefficient b^*_d.

By induction hypothesis, $H^*(T) = \lfloor S^*(T) \rfloor$. Furthermore, since F is semi-regular, we have the exact sequence

$$0 \longrightarrow K_{m,d} \longrightarrow R^*_d \xrightarrow{s_{m,d}} R^*_{d+d_m} \longrightarrow R_{d+d_m} \longrightarrow 0$$

where $R = \mathbb{K}[X_1, \ldots, X_n]/\langle F \rangle$ and $R^* = \mathbb{K}[X_1, \ldots, X_n]/\langle F^* \rangle$. As a consequence, for any $d \geq 0$, the coefficient b_d satisfies the recurrence relation:

$$b_{d+d_m} = b^*_{d+d_m} - b^*_d + \dim(K_{m,d})$$

where either $K_{m,d} = 0$ or $b_{d+d_m} = 0$. Since $s_{m,d}$ is defined from a space of dimension b^*_d to a space of dimension $b^*_{d+d_m}$, this can be rephrased as

$$b_d = \max\left(b^*_d - b^*_{d-d_m}, 0\right).$$

From Theorem 3.37 applied to $S(T)$, there exists δ such that

$$\forall d \in \{0, \ldots, \delta\},\ a_d = a^*_d - a^*_{d-d_m} > 0.$$

Furthermore, the induction hypothesis shows that there exists a degree δ^* such that

$$\forall d \in \{0, \ldots, \delta^*\},\ a^*_d = b^*_d > 0$$
$$\forall d > \delta^*, b^*_d = 0,$$

and that δ^* is defined as in Theorem 3.37. In particular, it implies that $\delta^* \geq \delta$.

We shall prove that the Hilbert series H of F is equal to S, truncated at degree δ. Let $d \geq 0$:

- if $0 \leq d \leq \delta \leq \delta^*$:

$$\begin{aligned} b_d &= b^*_d - b^*_{d_m} \text{ since } d \leq \delta \\ &= a^*_d - a^*_{d_m} \text{ since } d \leq \delta^* \\ &= a_d \end{aligned}$$

- if $\delta < d$:

$$\begin{aligned} b_d &= \max\left(b^*_d - b^*_{d-d_m}, 0\right) \\ &= 0 \text{ since } b^*_d = 0 \text{ and } b^*_{d-d_m} \geq 0 \end{aligned}$$

And since $a_{\delta+1} \leq 0$, this proves that

$$H(T) = \lfloor S(T) \rfloor.$$

□

3.4.3. W-degree of regularity of semi-regular sequences

Another consequence of Theorem 3.37 is an explicit value for the degree δ of the Hilbert series of an ideal defined by a semi-regular sequence with $m = n + 1$ polynomials in n variables. In the homogeneous case, it is known that this degree is bounded by

$$\delta = \min\left(\sum_{i=1}^{n} d_i - n, \left\lfloor \frac{\sum_{i=1}^{n+1} d_i - n}{2} \right\rfloor\right).$$

Proposition 3.39. *Let n be a positive integer, and $m = n + 1$. Let $W = (w_1, \ldots, w_n)$ be a system of weights, and $F = (f_1, \ldots, f_m)$ a system of W-homogeneous polynomials, and assume that the hypotheses of Theorem 3.38 are satisfied. For all $i \in \{1, \ldots, m\}$, let $d_i := \deg_W(f_i)$. Then the degree δ of the Hilbert series of $\langle F \rangle$ is given by:*

$$\delta = \min\left(\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} w_i, \left\lfloor \frac{\sum_{i=1}^{n+1} d_i - \sum_{i=1}^{n} w_i}{2} \right\rfloor\right).$$

Proof. Consider the system of weights $W^+ = (w_1, \ldots, w_n, 1)$, and the series S_{D,W^+} as defined in Theorem 3.37. It satisfies the hypotheses of Theorem 3.25, which implies that its coefficients are increasing up to degree

$$\sigma^+ = \min\left(\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} w_i, \left\lfloor \frac{\sum_{i=1}^{n+1} d_i - \sum_{i=1}^{n} w_i}{2} \right\rfloor\right).$$

Theorem 3.37 (statement ($\sigma 4$)) states that the degree δ of the Hilbert series of $\langle F \rangle$ satisfies

$$\delta = \sigma^+,$$

hence the result. □

In the end of this section, we show how the results from [Bar+05] and [Bar04, Chap. 4] about the degree of regularity of semi-regular homogeneous sequences can be adapted to the weighted case.

Theorem 3.40. *Let k and n be non-negative integers and let $m := n + k$. Let w_0 and d_0 be non-negative integers such that $w_0 \mid d_0$. Consider the system of n weights $W = (w_0, \ldots, w_0, 1)$ and the system of m W-degrees $D = (d_0, \ldots, d_0)$. Let $F = (f_1, \ldots, f_m) \subset A = \mathbb{K}[X_1, \ldots, X_n]$ be a semi-regular sequence of weighted homogeneous polynomials with W-degree D. Then the asymptotic expansion of the W-degree of regularity d_{reg} of F as n tends to infinity is given by*

$$d_{\text{reg}} = n\left(\frac{d_0 - w_0}{2}\right) - \alpha_k \sqrt{n\left(\frac{d_0^2 - w_0^2}{6}\right)} + O(n^{1/4}).$$

Remark 3.41. In the non-weighted case, this asymptotic expansion is

$$d_{\text{reg}} = n\left(\frac{d_0 - 1}{2}\right) - \alpha_k \sqrt{n\left(\frac{d_0^2 - 1}{6}\right)} + O(n^{1/4}).$$

Overall, the bound is improved by $O(nw_0) = O(\sum w_i)$.

Proof. Let $I \coloneqq \langle F \rangle$. The Hilbert series of A/I is given by

$$\mathrm{HS}_{A/I}(T) = \left\lfloor \frac{(1-T^{d_0})^m}{(1-T^{w_0})^{n-1}(1-T)} \right\rfloor.$$

Write

$$H(T) = \frac{(1-T^{d_0})^m}{(1-T^{w_0})^{n-1}(1-T)} = \sum_{d=0}^{\delta} a_d T^d;$$
$$H^*(T) = \frac{(1-T^{d_0/w_0})^{m-1}}{(1-T)^{n-1}} = \sum_{d=0}^{\delta} a_d^* T^d,$$

these series are related through

$$H(T) = H^*(T^{w_0}) \cdot \frac{1-T^{d_0}}{1-T} = H^*(T^{w_0}) \cdot \left(1 + T + \cdots + T^{d_0-1}\right).$$

It means for the coefficients that for any d in $\mathbb{N}$:

$$a_d = a^*_{\lfloor d/w_0 \rfloor} + \cdots + a^*_{\lceil (d-d_0+1)/w_0 \rceil}$$

If truncated before its first non-positive coefficient, the series H^* is the Hilbert series of a semi-regular $\mathbb{1}$-homogeneous sequence of $m-1$ polynomials in $n-1$ variables, with degree d_0/w_0. Let δ^* be the degree of this truncated series, so that δ^*+1 is an upper bound for its degree of regularity. Statement ($\delta 2$) of Theorem 3.37 states that:

$$\forall\, d \in \{\delta^*+1, \ldots, \delta^* + d_0/w_0\},\ a_d^* \leq 0.$$

Let $\delta_0 \coloneqq w_0\delta^* + d_0$, we have

$$\delta^* < \frac{\delta_0 - d_0 + 1}{w_0} \leq \delta^* + 1$$

and

$$\frac{\delta_0}{w_0} = \delta^* + \frac{d_0}{w_0},$$

and as a consequence

$$a_{\delta^0} = a^*_{\lfloor \delta_0/w_0 \rfloor} + \cdots + a^*_{\lceil (\delta_0-d_0+1)/w_0 \rceil} \leq 0.$$

In other words, the degree of regularity d_{reg} of F is bounded by

$$w_0\delta^* < d_{\mathrm{reg}} \leq \delta_0 = w_0\delta^* + d_0.$$

The degree δ^* is the degree of the Hilbert series of a homogeneous semi-regular sequence, and as such it follows the asymptotic estimates proved in [Bar04, Chap. 4]. For example in our

setting where k is an integer and $m = n + k$, the asymptotic expansion of δ^* when n tends to infinity is given by

$$\delta^* + 1 = n\,\frac{d_0/w_0 - 1}{2} - \alpha_k \sqrt{n\,\frac{(d_0/w_0)^2 - 1}{6}} + O(n^{1/4})$$

where α_k is the largest root of the k'th Hermite polynomial. [1]

As a consequence,

$$\begin{aligned} d_{\mathrm{reg}} &= w_0\delta^* + O(1) \\ &= w_0\left(n\,\frac{d_0/w_0 - 1}{2} - \alpha_k\sqrt{n\,\frac{(d_0/w_0)^2 - 1}{6}} + O(n^{1/4})\right) + O(1) \\ &= n\,\frac{d_0 - w_0}{2} - \alpha_k\sqrt{n\,\frac{d_0^2 - w_0^2}{6}} + O(n^{1/4}). \end{aligned}$$

□

3.4.4. Fröberg's conjecture

Fröberg's conjecture states that homogeneous semi-regular sequences are generic among sequences of fixed degree. The fact that semi-regularity is a Zariski-open condition is a known fact (the proof is the same as for regularity), so the conjecture states that for any system of degrees, there exists a semi-regular homogeneous sequence with these degrees.

This conjecture extends naturally to the weighted case. In this case, semi-regularity is still a Zariski-open condition.

We extend here one known result from the homogeneous case (see for example [RRR91]), stating that Fröberg's conjecture is true in characteristic 0 if $m = n + 1$. We follow the proof given in [RRR91].

Proposition 3.42. *Let $m = n + 1$, $W = (w_1, \ldots, w_n)$ a reverse chain-divisible system of weights, $D = (d_1, \ldots, d_n)$ a strongly W-compatible system of degrees, and d_{n+1} an integer divisible by w_1. Write $f_{n+1} = (X_1 + X_2^{w_1/w_2} + \cdots + X_n^{w_1})^{d_{n+1}/w_1}$.*

Then the sequence $F := (X_1^{d_1/w_1}, \ldots, X_n^{d_n/w_n}, f_{n+1})$ is semi-regular.

Lemma 3.43. *Let f be a polynomial such that*

$$f \cdot f_{n+1} = 0 \in A = \mathbb{K}[X_1, \ldots, X_n]/(X_1^{d_1/w_1}, \ldots, X_n^{d_n/w_n}).$$

Let $\delta = \sum_{i=1}^{n}(d_i - w_i)$, then we have

$$\deg_W(f) \geq \frac{\delta - d_{n+1} + 1}{2}.$$

[1] In [Bar04, Chap. 4], the remainder $O(n^{1/4})$ was written as $o(\sqrt{n})$. However, it appears that in the proof, this $o(\sqrt{n})$ is a rewriting of $\sqrt{n}\cdot O(\sqrt{\Delta z})$, where $\Delta z = O(1/\sqrt{n})$.

Proof. If the W-degree of f is 0, this means that $(X_1 + X_2^{w_1/w_2} + \cdots + X_n^{w_1})^{d_{n+1}/w_1} = 0$ in A. Assume that $\deg_W(f) < (\delta - d_{n+1} + 1)/2$, that means that $\delta - d_{n+1} + 1 \geq 1$, so $\delta \geq d_{n+1}$. Consider the expansion of f_{n+1}. All its coefficients are nonzero since the base field has characteristic 0. Its support is the set of monomials of degree d_{n+1}. Since $d_{n+1} \leq \delta$, $\dim(\mathbb{K}[\mathbf{X}]/\langle f_1, \ldots, f_n\rangle)_{d_{n+1}} > 0$, which means that there exists at least one monomial with W-degree d_{n+1} which does not lie in the initial ideal of $\langle f_1, \ldots, f_n\rangle$. As a consequence, f is non-zero in the quotient.

Now assume that $\deg_W(f) > 0$. Write $B = \mathbb{K}[X_2, \ldots, X_n]/(X_2^{d_2/w_2}, \ldots, X_n^{d_n/w_n})$, $X = X_1$, $R = B[X]$, $d = d_1/w_1$, such that $A = R/X^d$. Let $S = (X + X_2^{w_1/w_2} + \cdots + X_n^{w_1})$, and let F be a weighted homogeneous polynomial in R with image f in A. By assumption, there exists $G \in R$ such that $S^{d_{n+1}/w_1} \cdot F = G \cdot X^d$. Differentiate this equality along X to obtain:

$$mS'S^{d_{n+1}/w_1-1}F + S^{(d_{n+1})/w_1}F = dG'X^{d-1} + G'X^d$$

which gives, modulo X^{d-1}

$$\begin{aligned} S^{d_{n+1}-1}(mF + SF') \equiv 0\,[X]^{d-1} &\implies S^{d_{n+1}/w_1}(mF + SF') \equiv 0\,[X]^{d-1} \\ &\implies S^{d_{n+1}/w_1+1}F' \equiv mFS^{d_{n+1}/w_1} \equiv 0\,[X]^{d-1} \end{aligned}$$

Since $X = X_1$ has weight w_1, F' is W-homogeneous with W-degree $\deg_W(f) - w_1$, and we can use the induction hypothesis on $F'\,[X] \in A$ and $\deg(F) = d_{n+1} + w_1$ to deduce:

$$\begin{aligned} \deg_W(f) = \deg_W(F) &= \deg_W(F') + 1 \\ &\geq \frac{(\delta - 1) - (d_{n+1} + 1) + 1}{2} + 1 \\ &\geq \frac{\delta - d_{n+1} + 1}{2}. \end{aligned}$$

□

Proof of the proposition. The proof given in [RRR91, before prop. 7] still holds in the weighted case. □

3.5. Algorithms and complexity

3.5.1. Critical pairs algorithms

As said previously, we can compute a Gröbner basis of a weighted homogeneous ideal generated by a system F by running usual algorithms on $\hom_W(F)$. The following proposition shows that this is correct:

Proposition 3.44. *Let $F = (f_1, \ldots, f_m)$ be a family of polynomials in $\mathbb{K}[X_1, \ldots, X_n]$, assumed to be weighted homogeneous for a system of weights $W = (w_1, \ldots, w_n)$. Let $<_1$ be a monomial order, G the reduced Gröbner basis of $\hom_W(F)$ for this order, and $<_2$ the pullback of $<_1$ through $\hom_W$. Then*

1. *all elements of G are in the image of $\hom_W$;*
2. *the family $G' := \hom_W^{-1}(G)$ is a reduced Gröbner basis of the system F for the order $<_2$.*

Proof. The morphism $\hom_W$ preserves S-polynomials, in the sense that

$$S\text{-pol}(\hom_W(f), \hom_W(g)) = \hom_W(S\text{-pol}(f, g)).$$

Recall that we can compute a reduced Gröbner basis by running the Buchberger algorithm, which involves only multiplications, additions, tests of divisibility and S-polynomials. Since all these operations are compatible with $\hom_W$, if we run the Buchberger algorithm on both F and $\hom_W(F)$ simultaneously, they will follow exactly the same computations up to application of $\hom_W$. The consequences on the final reduced Gröbner basis follow. □

This shows that all pairs-based algorithms (for example Buchberger, F_4 and F_5) can be made to take into account a weighted homogeneous structure, simply by wrapping $\hom_W$ around the algorithm.

3.5.2. Adapting the Matrix-F_5 algorithm

In order to study the complexity of algorithm F_5, we want to run algorithm Matrix-F_5 on a weighted homogeneous system. Recall that this algorithm works by constructing a matrix of linear bases of I_d incrementally, and does not use critical pairs to guide the reductions.

If we wrap the algorithm with $\hom_W$, Proposition 3.44 applies and shows that the result is correct. However, it also states that we can compute a Gröbner basis of $\langle \hom_W(F) \rangle$ by working with S-polynomials in the image of $\hom_W$, so those rows whose label monomial does not lie in this image will be effectively useless.

There are two equivalent ways around this problem. We can run the algorithm on $\hom_W(F)$, only considering monomials in the image of $\hom_W$ for the labels. Or we can adapt Matrix-F_5 to use the weighted degree instead of the total degree. We present the second solution.

The only real difference is at line 14: if μ has W-degree $d - d_i$, μ/x_j has W-degree $d - d_i - w_j$ and so the polynomial f comes from a row of the matrix at W-degree $d - d_i - w_j + d_i = d - w_j$.

Theorem 3.45. *Let $F = (f_1, \dots, f_m)$ be a system of W-homogeneous polynomials, and let $P = \prod_{i=1}^{n} w_i$. Then Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a $d_{\max}$-Gröbner basis of $\langle F \rangle$ in time*

$$O\left(d_{\text{reg}} \frac{1}{P^3} \binom{n + d_{\max} - 1}{d_{\max}}^3\right). \tag{3.14}$$

If $d_{\max} \geq d_{\text{reg}}$ and the monomial order is W-GRevLex, then Matrix-$\mathsf{F}_5(F, d_{\max})$ computes a Gröbner basis of F. If F is in Noether position with respect to the variables $X_{m+1}, \dots, X_n$, then d_{reg} is bounded by

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - w_i) + \max\{w_j\}.$$

If F is in simultaneous Noether position with respect to the order $X_1 > X_2 > \cdots > X_m$, then d_{reg} is bounded by

$$d_{\text{reg}} \leq \sum_{i=1}^{m} (d_i - w_i) + w_m.$$

Algorithm 3.1 Matrix-F_5 (weighted case)

Input: $G \subset \mathbb{K}[\mathbf{X}]$ a truncated Gröbner basis of $\langle F \rangle$ at W-degree $d_{\max}$
Output: $F \subset \mathbb{K}[\mathbf{X}]$, $d_{\max} \in \mathbb{N}$ an implicit monomial ordering, $W \in \mathbb{N}^n$

$G \leftarrow \varnothing$
for $d = 0$ **to** $d_{\max}$ **do**
  $N_d \leftarrow$ number of monomials of W-degree d
  $M_{d,0}, \tilde{M}_{d,0} \leftarrow$ matrix with o rows and N_d columns
  for $i = 1$ **to** m **do**
    $M_{d,i} \leftarrow M_{d,i-1}$
    if $d = d_i$ **then**
      $M_{d,i} \leftarrow M_{d,i} \cup$ row f_i with signature $(i, 1)$
    else if $d > d_i$ **then**
      for all μ monomial of W-degree $d - d_i$ **do**
        if μ is not the leading term of a row in $\tilde{M}_{d-d_i,i-1}$ **then**
          $j \leftarrow$ largest j such that X_j divides mu
          $\mu' \leftarrow \mu / X_j$
          $\tilde{f} \leftarrow$ row of $\tilde{M}_{d-w_j,i}$ with signature (i, μ')
          $M_{d,i} \leftarrow M_{d,i} \cup$ row $X_j \tilde{f}$ with signature (i, μ)
        end if
      end for
    end if
  end for
  $\tilde{M}_{d,i} \leftarrow \mathsf{EchelonForm}(M_{d,i})$
  $R \leftarrow$ number of rows in $M_{d,i}$
  $G \leftarrow G \cup \{\mathrm{row}(\tilde{M}_{d,i}, j) \mid j \in \{1, \ldots, R\}, \mathrm{row}(\tilde{M}_{d,i}, j) \neq \mathrm{row}(M_{d,i}, j)\}$
end for
return G

Proof. The complexity is computed as in the homogeneous case, in terms of the maximal degree $d_{\max}$ and the size of the matrices. The size of the matrices depends on the number of monomials at a given W-degree, and Proposition 1.19 shows that asymptotically, this number of monomials is divided by P. The bounds on d_{reg} come respectively from Theorem 3.27 and Theorem 3.28. □

3.5.3. Thin-grained complexity of Matrix-F$_5$

As in the homogeneous case, we wish to obtain sharper complexity bounds. We will show that the computations of Section 2.5.3 can be carried in a weighted setting.

Let $W = (w_1, \dots, w_n)$ be a system of weights, and $(f_1, \dots, f_m) \in \mathbb{K}[\mathbf{X}]$ a system of weighted homogeneous polynomials in simultaneous Noether position with respect to the order $X_1 > \cdots > X_n$. Let $d_1, \dots, d_m$ be the respective W-degrees of the polynomials $f_1, \dots, f_m$.

We also denote by:

- $A_i = \mathbb{K}[X_1, \dots, X_i]$, and $A = A_n$;
- $S_i = \sum_{j=2}^{i} w_i \frac{\gcd(w_1, \dots, w_j)}{\gcd(w_1, \dots, w_{j-1})}$ (see the definition of A in Prop. 1.20), and $S = S_n$;
- $P_i = \prod_{j=1}^{i} w_j$, and $P = P_n$;
- $I_i = \langle f_1, \dots, f_i \rangle$, and $I = I_m$;
- $\hom_W(f_j) = \hom_W(f_j)$;
- $\hom_W(I_i) = \langle \hom_W(f_1), \dots, \hom_W(f_i) \rangle$, and $\hom_W(I) = \hom_W(I_m)$;
- $D_i = \prod_{j=1}^{i} \frac{d_j}{w_j}$ (with the convention $D_0 = 0$);
- $\tilde{D}_i = \deg(\hom_W(I_i)) = \prod_{j=1}^{i} d_j$;
- $d_{\text{reg}}^{(i)}$ the degree of regularity of F_i (or of $\hom_W(F_i)$) ;
- G_i the W-GRevLex Gröbner basis of I_i as given by Matrix-F$_5$.

With these notation and hypotheses, we are going to prove the following theorem:

Theorem 3.46. *The complexity of weighted homogeneous Matrix-F$_5$ algorithm (Algorithm 3.1) is:*

$$\begin{aligned} C_{\mathsf{F}_5} &= O\left(\sum_{i=2}^{m} (D_{i-1} - D_{i-2}) M_{d_{\text{reg}}^{(i)}, W}(i) M_{d_{\text{reg}}^{(i)}, W}(n) \right) \\ &= O\left(\sum_{i=2}^{m} \frac{D_{i-1} - D_{i-2}}{P_i P_n} \binom{i + d_{\text{reg}}^{(i)} - 1}{d_{\text{reg}}^{(i)}} \binom{n + d_{\text{reg}}^{(i)} - 1}{d_{\text{reg}}^{(i)}} \right) \end{aligned}$$

As in the homogeneous case, this theorem relies on the structure lemma:

Lemma 3.47 (Structure lemma, weighted case)**.** *Let $i \in \{1, \dots, m\}$, and $g \in G_i$ with signature (j, μ). Then:*

- $j \leq i$
- $\mathsf{LT}(g) \in \mathbb{K}[X_1, \ldots, X_i]$
- $\mu \in \mathbb{K}[X_1, \ldots, X_{i-1}]$

Proof. The proof of the homogeneous version of the lemma (Lemma 2.51) does not depend on the graduation. □

With the structure lemma, we can count the number of polynomials in each intermediate basis, W-degree by W-degree:

Proposition 3.48. *Let $(f_1, \ldots, f_m)$ be a W-homogeneous system in simultaneous Noether position with respect to the order $X_1 > \cdots > X_n$. Let G_i be a reduced Gröbner basis of $(f_1, \ldots, f_i)$ for the W-GRevLex monomial ordering, for $1 \leq i \leq m$. Then the number of polynomials of W-degree d in G_i whose leading term does not belong to $\mathsf{LT}(G_{i-1})$ is bounded by $b_{d,i}$, defined by the generating series*

$$B_{W,i}(T) = \sum_{d=0}^{\infty} b_d^{(W,i)} T^d = T^{d_i} \prod_{k=1}^{i-1} \frac{1 - T^{d_k}}{1 - T^{w_k}}. \tag{3.15}$$

Proof. The proof of Proposition 2.52 still holds in a weighted setting, substituting formula (3.18) for the Hilbert series of a weighted homogeneous regular sequence. □

As in Section 2.1.3, we define for any $l \leq n$ the endomorphism θ_l of $\mathbb{K}[\mathbf{X}]$, sending X_i on itself for $i \leq l$ and on 0 otherwise. Note that by Proposition 1.44, $\theta_l(F)$ is still in Noether position for all $l \in \{1, \ldots, m\}$.

Theorem 3.49. *The complexity of weighted homogeneous Matrix-F_5 algorithm (Algorithm 3.1) is:*

$$\begin{aligned} C_{\mathsf{F}_5} &= O\left(\sum_{i=2}^{m} \sum_{d=0}^{\infty} b_{d+d_i}^{(W,i)} N_{W_i, d+d_i} N_{W, d+d_i} \right) \\ &= O\left(\sum_{i=2}^{m} \sum_{d=0}^{\infty} \frac{b_{d+d_i}^{(W,i)}}{P_i P_n} \binom{d + d_i + S_i}{d + d_i + S_i - i + 1} \binom{d + d_i + S_i}{d + d_i + S_i - n + 1} \right) \end{aligned} \tag{3.16}$$

Proof of Theorem 3.49. The proof of Theorem 2.54 adapts to a weighted setting: $b_{d+d_i}^{(W,i)}$ bounds the number of polynomials reduced with signature $(i, \bullet)$ and W-degree d, $N_{W_i,d}$ bounds the number of rows used for each reduction at degree d for a polynomial with signature $(i, \bullet)$, and $N_{W,d}$ bounds the length of the rows. The second equality follows from the bounds of Proposition 1.20 on $N_{W_i,d}$. □

Proof of Theorem 3.46. With the same notations, for any $i \in \{1, \ldots, m\}$, $B_{W,i}(1)$ represents the number of reduced polynomials during the computation of a Gröbner basis of $(f_1, \ldots, f_i)$. Equation (3.15) states that $B_{W,i}(T)/T^{d_i}$ is the Hilbert series of $(\theta_i(f_1), \ldots, \theta_i(f_{i-1}))$, and so its value at $T = 1$ is the degree of that ideal, or D_{i-1}. Therefore we know that the number of reduced polynomials with label (μ, f_i) is bounded by $D_{i-1} - D_{i-2}$.

The rest of the factors are obtained in the same way as in the proof of Theorem 3.49, using the fact that all polynomials with label $(i, \bullet)$ are reduced at a W-degree bounded by $d_{\mathrm{reg}}^{(i)}$. □

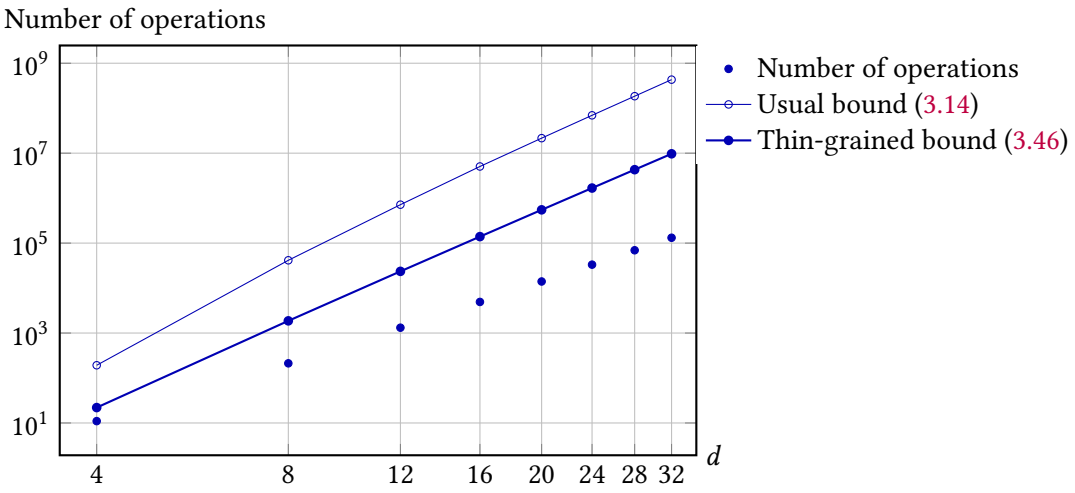

Figure 3.7: Number of operations needed by Matrix-F_5 on a generic $[4, 2, 1]$-homogeneous system with degree $[d, d, d]$

The behavior of both bounds is shown on Figure 3.7.

3.5.4. FGLM and computational strategy for zero-dimensional systems

Let $F = (f_1, \ldots, f_n) \subset \mathbb{K}[X_1, \ldots, X_n]$ be a zero-dimensional affine system. In this situation, we usually want to enumerate all solutions of the system. As described in Section 2.3.1, one can generically obtain a triangular description of the solutions from a lexicographical basis of the ideal, and it can be computed from another Gröbner basis using algorithm FGLM.

Recall from Section 2.5.4 that the complexity of algorithm FGLM mainly depends on the degree of the ideal. The weighted version of Bézout's bound implies that at least generically, applying $\hom_W$ to a zero-dimensional ideal multiplies the number of solutions by the product of the weights. So wrapping $\hom_W$ around calls to FGLM incurs an additional complexity factor of $\prod w_i$.

On the other hand, algorithm FGLM does not need adapting in order to work with systems with a weighted homogeneous structure: it takes as input a Gröbner basis for any order, in particular one can use a W-GRevLex basis.

All in all, Bézout's bound gives the following complexity for FGLM for a weighted system:

Theorem 3.50. *The complexity of algorithm FGLM for computing the Gröbner basis of a zero-dimensional ideal defined by $(f_1, \ldots, f_n)$, polynomials with a W-homogeneous structure, and with respective W-degrees $(d_1, \ldots, d_n)$.*

$$O\left(n\left(\frac{\prod_{i=1}^{n} d_i}{\prod_{i=1}^{n} w_i}\right)^3\right). \tag{3.17}$$

Experimentally, the improvement of [Fau+14] appears to work for bases for the W-GRevLex order, so that ω can be taken as the complexity exponent.

3.5.5. Other algorithms

The gain from the reduced number of monomials applies to other algorithms as well, provided they are run on $\hom_W(F)$ if they are only using critical pairs, or use the W-GRevLex order otherwise.

For example, when computing elimination bases for positive dimensional systems, depending on the situation, it may be interesting to perform a two-steps computation, or to simply use one of the direct algorithms with the desired order. In the former case, FGLM cannot be used for the change of ordering. A common choice in this situation is the Gröbner walk. This algorithm is much more complex and to the best of our knowledge, does not have good complexity estimates. However, it involves computing successive Gröbner bases using algorithm F_4 or F_5 as a blackbox. As such, assigning weights to a polynomial system will yield similar improvements for the computing time.

3.6. Experiments

In this section, we present some applications where taking into account the weighted structure of the system yields speed-ups. For each system, we compare two strategies: the "standard" strategy consists of computing a Gröbner basis without considering the weighted structure; the "weighted" strategy is the strategy we described in Section 3.5. For all these examples, we use a more compact notation for degrees and weights, so that for example, $(2^3, 1)$ is equivalent to $(2, 2, 2, 1)$.

3.6.1. Generic systems

First, we present some timings obtained with generic systems, in the complete intersection ($m = n$), the positive-dimensional ($m < n$) and the over-determined ($m > n$) cases. In all cases, we fix a system of weights $W = (w_1, \ldots, w_n)$ and a system of W-degrees $D = (d_1, \ldots, d_m)$, and we pick at random m polynomials $(f_i)_{i=1\ldots m}$, such that for any $i \in \{1, \ldots, m\}$, f_i has dense support in the set of monomials with W-degree less than or equal to d_i.

For zero-dimensional regular sequences, we compute a lexicographical Gröbner basis using a two-steps strategy in Magma, with algorithm F_4 as a first step (first block of lines in Table 3.2a) and algorithm FGLM for the change of ordering (Table 3.2b).

For over-determined systems, we compute a Gröbner basis for the GRevLex ordering, using algorithm F_4 from Magma (second block of lines in Table 3.2a).

For positive-dimensional systems, we compute a basis for an elimination order, using a two-steps strategy with FGb[2]: first we compute a GRevLex basis with algorithm F_4 (third block of lines in Table 3.2a), and then we compute a basis for the wanted elimination order, again with F_4 (Table 3.2c). In this table, the second column describes what variables we eliminate:

[2] The Gröbner basis algorithms from Magma seem to behave strangely with elimination orders, as seen in the detailed logs, and it coincides with significant slowdowns. This behavior was not observed on other implementations of the same algorithms: F_4 from FGb and Buchberger from Singular [Singular]. For example, for the system in the first line of table 3.2c, without the weights, with Magma's F_4 algorithm, the first degree fall comes at step 4, and the algorithm needs more than 66 steps to compute the basis. With FGb's implementation of F_4 in Maple, the first degree fall appears at step 13, and the algorithm finishes at step 32.

for example, 3 means that we eliminate the first 3 variables, while $1 \to 3$ means that we first eliminate the first variable, then the next 2 variables, again resulting in a basis eliminating the first 3 variables.

For algorithm F_4 with the GRevLex ordering, the behavior we observe is coherent with the previous complexity studies: we observe some speed-ups when taking into account the weighted structure of the system, and these speed-ups seem to increase with the weights. However, the speed-ups cannot be expected to match rigorously the ones predicted by the complexity bounds, because the systems are usually not regular for the standard strategy. Experiments also confirm that it is more effective to order the variables with highest weight first.

For the lexicographical ordering with FGLM, we also observe some speed-ups when applying the weights (we will observe this behavior again in Section 3.6.2). These differences are not explained by the theoretical complexity bounds, since both ideals have the same degree in each case. However, it appears that the slower FGLM runs are those where the FGLM matrix is denser, and that this difference in density seems to match quantitatively the speed-ups we observe.

Finally, for elimination bases, the results are similar to what we observed with the GRevLex ordering: when possible, one should take into account the weights, and order the variables such that the smallest weights are also the smallest variables. However, when eliminating variables, the largest variables need to be the ones that should be eliminated. If the variables need to be ordered such that those with the smallest weights are first, in most cases, taking into account the weighted structure is still profitable. However, if the smallest weight is on the largest variable and there is only one such variable, this is no longer true (see for example the second line in Table 3.2c). Experiments suggest that these systems naturally possess a good weighted structure for the weights $(1, \dots, 1)$: their construction ensures that every such polynomial of total degree d will have a large homogeneous component at degree $d/2$, and the higher degree components will be small, and divisible by large powers of X_1. On the other hand, with weights $(1, 2, \dots, 2)$, the same polynomial will have a large W-homogeneous component at W-degree d, overall leading to reductions at higher degree (an example is given in Table 3.3).

We conclude this section with timings illustrating the consequences of the estimates of the degree of regularity of a system, depending on the order of the variables (Section 3.3.2). For this purpose, we generate a generic system of W-degree (60^4) with weights $(20, 5, 5, 1)$. Then we compute a W-GRevLex Gröbner basis for the orders $X_1 > \cdots > X_4$ (smallest weights last) and for the reverse order $X_4 < \cdots < X_1$. We give the degree of regularity, the value predicted by the previous bound (3.9), by the new bound (3.10) and by the conjectured bound (3.12), as well as the timings. This experiment was run using algorithm F_5 from the FGb library, the results are in Table 3.4.

3.6.2. Discrete Logarithm Problem

Taking advantage of a weighted homogeneous structure has allowed the authors of the article [Fau+13] to obtain significant speed-ups for solving a system arising from the DLP on Edwards elliptic curves [Gau09]. They observed that the system of equations they had to solve has symmetries, and rewrote it in terms of the invariants of the symmetry group. For a system

Table 3.2: Benchmarks with Magma for generic systems

Parameters	Without weights (s)	With weights (s)	Speed-up
$n = 8, W = (2^6, 1^2), D = (4^8)$	8.0	2.5	3.2
$n = 9, W = (2^7, 1^2), D = (4^9)$	101.2	12.5	8.1
$n = 7, W = (2^5, 1^2), D = (8^{15})$	31.6	7.5	4.2
$n = 7, W = (2^5, 1^2), D = (8^{14})$	29.0	9.4	3.1
$n = 7, W = (2^5, 1^2), D = (8^{13})$	40.0	12.0	3.3
$n = 5, m = 4, W = (2^4, 1), D = (8^4)$	2.6	0.2	13.0
$n = 5, m = 4, W = (1, 2^4), D = (8^4)$	2.5	0.3	8.3
$n = 5, m = 4, W = (1^3, 2^2), D = (4^4)$	23.6	0.0	2360.0
$n = 5, m = 4, W = (2^2, 1^3), D = (4^4)$	407.5	0.0	40 750.0

(a) Benchmarks for the F_4 algorithm for the GRevLex ordering

Parameters	Degree	Without weights (s)	With weights (s)	Speed-up
$n = 8, W = (2^6, 1^2), D = (4^8)$	1024.0	500.4	495.0	1.0
$n = 9, W = (2^7, 1^2), D = (4^9)$	2048.0	11 995.8	7462.1	1.6

(b) Benchmarks for the FGLM algorithm (lexicographical ordering)

Parameters	Elim. vars.	Without weights (s)	With weights (s)	Speed-up
$n = 5, m = 4, W = (2^4, 1), D = (8^4)$	1	120.3	12.0	10.0
$n = 5, m = 4, W = (1, 2^4), D = (8^4)$	1	27.6	30.4	0.9
$n = 5, m = 4, W = (1^3, 2^2), D = (4^4)$	2	146.3	6.9	21.2
$n = 5, m = 4, W = (1^3, 2^2), D = (4^4)$	$1 \longrightarrow 2$	162.0	3.3	49.1
$n = 5, m = 4, W = (2^2, 1^3), D = (4^4)$	1	>750	0.1	>7500
$n = 5, m = 4, W = (2^2, 1^3), D = (4^4)$	$1 \longrightarrow 2$	NA	0.1	NA
$n = 5, m = 4, W = (2^2, 1^3), D = (4^4)$	$1 \longrightarrow 2 \longrightarrow 3$	NA	7.9	NA

(c) Benchmarks for the F_4 algorithm for elimination

Table 3.3: Size of the W-homogeneous components for a generic polynomial with W_0-degree 4 for $W_0 = (1, 2, 2, 2)$

W-degree	$W = (1, 2, 2, 2)$	$W = (1, 1, 2, 2)$	$W = (1, 1, 1, 1)$
0	1	1	1
1	1	2	4
2	4	5	10
3	4	6	4
4	10	6	1

Table 3.4: Impact of the order of the variables on the degree of regularity and the computation times (generic weighted homogeneous system with W-degree (60^4) w.r.t. $W = (20, 5, 5, 1)$)

Order	d_{reg}	Macaulay's bound (3.9)	Bound (3.15)	Conjectured bound (3.12)	F_5 time
$X_1 > X_2 > X_3 > X_4$	210	229	210	210	101.9
$X_4 > X_3 > X_2 > X_1$	220	229	229	220	255.5

Table 3.5: Benchmarks with FGb and Magma for DLP systems

System	deg(I)	F_5 w (s)	F_5 std (s)	Speed-up for F_5	FGLM w (s)	FGLM std (s)	Speed-up for FGLM
DLP Edwards: $n = 4$, $W = (2^3, 1)$, $D = (8^4)$	512	0.1	0.1	1.0	0.1	0.1	1.0
DLP Edwards: $n = 5$, $W = (2^4, 1)$, $D = (16^5)$	65 536	935.4	6461.2	6.9	2164.4	6935.6	3.2

(a) Benchmarks with FGb

System	deg(I)	F_4 w (s)	F_4 std (s)	Speed-up for F_4	FGLM w (s)	FGLM std (s)	Speed-up for FGLM
DLP Edwards: $n = 4$, $W = (2^3, 1)$, $D = (8^4)$	512	1	1	1.0	1	27	27
DLP Edwards: $n = 5$, $W = (2^4, 1)$, $D = (16^5)$	65 536	6044	56 105	9.3	∞	∞	NA

(b) Benchmarks with Magma

in n equations, these invariants are

$$\begin{aligned} E_1 &= e_1(X_1^2, \ldots, X_n^2) \\ E_2 &= e_2(X_1^2, \ldots, X_n^2) \\ &\vdots \\ E_{n-1} &= e_{n-1}(X_1^2, \ldots, X_n^2) \\ E_n &= e_n(X_1, \ldots, X_n). \end{aligned}$$

The system they obtained is sparser, but does not have a good homogeneous structure. In particular, the highest total degree components of the system do not form a regular sequence, and in practice, a Gröbner basis computation will follow many degree falls.

However, the system had a weighted homogeneous structure for the weights $(2, \ldots, 2, 1)$ (only E_n has weight 1), with respective W-degree $(2^n, \ldots, 2^n)$. The highest W-degree components forming a sequence in simultaneous Noether position with respect to the order $E_1 > E_2 > \cdots > E_n$, one could compute a Gröbner basis without any W-degree fall, with complexity bounded by the estimates (3.14) and (3.17).

This system also illustrates the role of the choice of the system of weights: recall that we

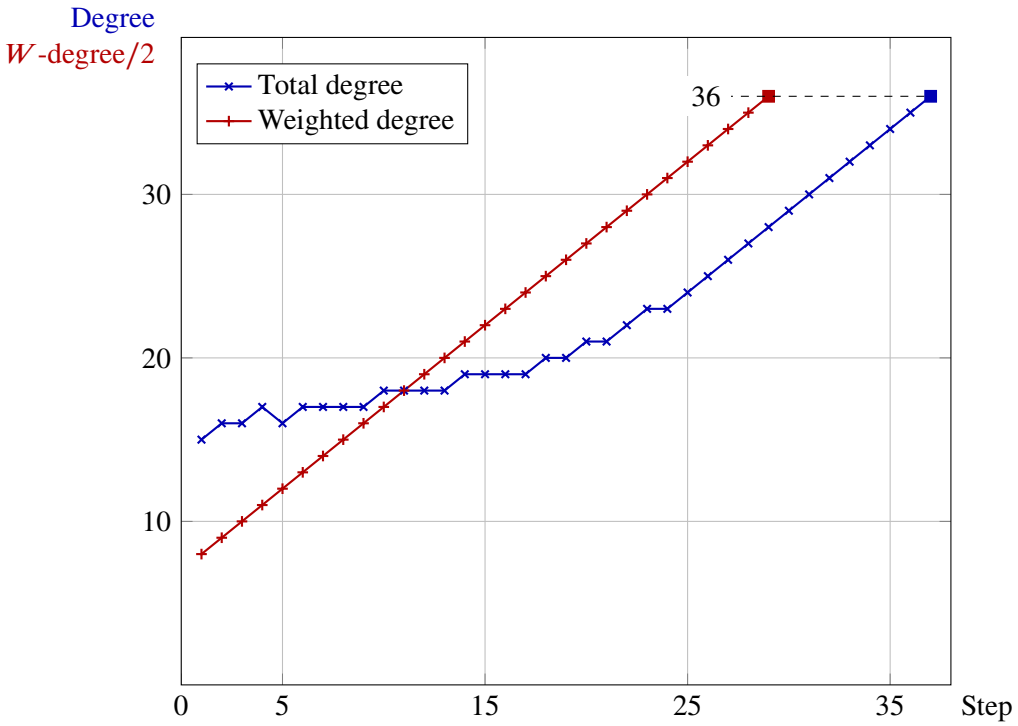

Figure 3.8: Degree at each reduction step in a run of Algorithm F_5 on a DLP system ($n = 5$, $W = (2^3, 1)$, $D = (16^5)$)

want to avoid degree falls, that is reductions to zero of the highest W-degree components of the polynomials.

To compare the behavior of the algorithm with the total degree and the appropriate W-degree, we plotted the degree for each reduction step in a run of algorithm F_4 in Figure 3.6.

On the blue curve, any horizontal or downwards segment is a step where the degree is less than or equal to the degree of the previous step: a degree fall. Indeed, when using the total degree, the system was not regular in the affine sense, and degree falls are to be expected. On the other hand, the red curve shows the progress of the algorithm when using the weights. It is an example of a regular run: there is no degree fall.

In this example, the final W-degree reached with the weights is exactly twice the final degree reached without the weights, so that the matrices built and reduced have approximately the same size. But when there are degree falls, evaluating the highest degree reached is not sufficient for counting the number of steps.

3.6.3. Polynomial inversion

The polynomial inversion problem consists in finding polynomial relations between polynomials. More precisely, given a system of polynomial equations

$$\begin{cases} f_1(X_1,\dots,X_n) &= 0 \\ f_2(X_1,\dots,X_n) &= 0 \\ & \vdots \\ f_m(X_1,\dots,X_n) &= 0\,, \end{cases}$$

we want to compute all the relations of the form

$$g_i(f_1,\dots,f_r) = 0.$$

One can compute these relations with Gröbner bases by computing an elimination ideal: consider the ideal generated by the polynomials

$$\begin{array}{rcl} T_1 & - & f_1(X_1,\dots,X_n) \\ T_2 & - & f_1(X_1,\dots,X_n) \\ & \vdots & \\ T_m & - & f_m(X_1,\dots,X_n) \end{array}$$

in $R \coloneqq \mathbb{K}[X_1,\dots,X_n,T_1,\dots,T_m]$. Order R with an elimination order on the variables $X_1,\dots,X_n$, recall that it is an order such that

$$m_X(X_1,\dots,X_n)m_T(T_1,\dots,T_m) <_{\text{elim}} m'_X(X_1,\dots,X_n)m'_T(T_1,\dots,T_m) \tag{3.18}$$

$$\iff \begin{cases} m_X <_X m'_X \\ \text{or} \\ m_X = m'_X \text{ and } m_T <_T m'_T \end{cases}$$

for some monomial orders $<_X$ and $<_T$. The usual choice is the block-GRevLex order $\textsc{Elim}_n$.

This problem can benefit from being given a weighted structure [Tra96, sec. 6.1]. For any $i \in \{1,\dots,m\}$, let d_i be the degree of f_i. By setting the weight of T_i to be d_i, the monomial T_i becomes part of the highest W-degree component of $T_i - f_i(X_1,\dots,X_n)$, giving this equation a weighted homogeneous structure.

More precisely:

Proposition 3.51. *Let $f_1,\dots,f_m$ be a system of polynomials with respective degree $d_1,\dots,d_m$ in $\mathbb{K}[X_1,\dots,X_n]$. Consider the algebra $R \coloneqq \mathbb{K}[X_1,\dots,X_n,T_1,\dots,T_m]$, graded with the weights $W = (1,\dots,1,d_1,\dots,d_m)$, and consider the system $F = (T_1 - f_1(\mathbf{X}),\dots,T_m - f_m(\mathbf{X}))$ in R. Then the system F^h formed with the highest W-degree components of F is in Noether position with respect to the variables $T_1,\dots,T_m$, and in particular it forms a regular sequence.*

Proof. By the choice of the weights, the system F^h is defined by

$$F^h = (T_1 - f_1^h(\mathbf{X}),\dots,T_m - f_m^h(\mathbf{X})),$$

Table 3.6: Benchmarks with Magma on some polynomial inversion systems

System	Without weights (s)	With weights (s)	Speed-up
Cyclic invariants, $n = 4$	4.2	0.0	140.0
Cyclic invariants, $n = 5, k = 12$	2612.6	54.7	47.8
Cyclic invariants, $n = 5$	$> 75\,000$ [a]	392.7	NA
Cyclic invariants, $n = 6, k = 14$	32 987.6	2787.7	11.8
Cyclic invariants, $n = 6, k = 15$	$>280\,000$ [a]	14 535.4	NA
Dihedral invariants, $n = 5$	$> 70\,000$ [a]	6.3	NA
Generic monomials, $d = 2, n = 24, m = 48$	216.1	0.2	1350.6
Generic monomials, $d = 2, n = 25, m = 50$	14 034.7	0.1	116 955.8
Generic monomials, $d = 2, n = 26, m = 52$	14 630.6	0.2	66 502.7
Generic monomials, $d = 2, n = 27, m = 54$	8887.6	0.2	55 547.5
Generic monomials, $d = 3, n = 11, m = 22$	370.9	0.1	6181.7
Generic monomials, $d = 3, n = 12, m = 24$	4485.0	0.2	26 382.4
Matrix minors, $n = 5, 7 \times 7, r = 3$	125.7	93.3	1.3
Matrix minors, $n = 6, 7 \times 7, r = 3$	1941.0	1029.1	1.9
Matrix minors, $n = 6, 8 \times 8, r = 3$	4115.8	2295.8	1.8
Matrix minors, $n = 4, 6 \times 6, r = 5$	612.6	159.2	3.8
Matrix minors, $n = 4, 7 \times 7, r = 6$	8043.3	2126.9	3.8
Matrix minors, $n = 4, 7 \times 10, r = 7$	69 386.1	43 910.1	1.6

a. Memory usage was over 120 GB

(a) First step (F_4 for the GRevLex order)

where for any $i \in \{1, \ldots, m\}$, f_i^h is the highest degree component of f_i. As a consequence, by the characterization 4 of the Noether position, the system F^h is indeed in Noether position with respect to the variables $T_1, \ldots, T_m$. □

In Tables 3.6, we present timings for a few systems with this kind of problem:

- group invariants [Stu08]: given a group, compute its fundamental invariants, and then the relations between these invariants. Since these examples can lead to very long computations, in some cases, we only compute the relations between the k first invariants;
- monomials: given m monomials of degree d in $\mathbb{K}[X_1, \ldots, X_n]$, compute the relations between them;
- matrix minors: given a $p \times q$ matrix of linear forms in n indeterminates, compute all its minors of rank r as polynomials in the $X_{i,j}$'s, and compute the relations between them.

In each case, we compute an elimination basis using a two-steps strategy: first we compute a GRevLex basis (Table 3.6a), then we compute the elimination basis (Table 3.6b). In Table 3.6c, we show some timings for the computation of the elimination basis directly from the input system. All these experiments were run using algorithm F_4 from Magma.

As in the previous section, we plot in Figure 3.9 the degree of each reduction step on a run of Algorithm F_4 for the GRevLex step in finding the relations between a random set of monomials.

System	Without weights (s)	With weights (s)	Speed-up
Cyclic invariants, $n = 4$	7.0	0.1	70.0
Cyclic invariants, $n = 5$, $k = 12$	1683.2	70.7	23.8
Cyclic invariants, $n = 5$	NA	382.5	NA
Cyclic invariants, $n = 6$, $k = 14$	9236.4	1456.0	6.3
Cyclic invariants, $n = 6$, $k = 15$	NA	7179.7	NA
Dihedral invariants, $n = 5$	NA	20.3	NA
Generic monomials, $d = 2$, $n = 24$, $m = 48$	250.3	117.4	2.1
Generic monomials, $d = 2$, $n = 25$, $m = 50$	13 471.2	15 932.9	0.8
Generic monomials, $d = 2$, $n = 26$, $m = 52$	17 599.5	8054.2	2.2
Generic monomials, $d = 2$, $n = 27$, $m = 54$	9681.0	3605.6	2.7
Generic monomials, $d = 3$, $n = 11$, $m = 22$	624.5	199.9	3.1
Generic monomials, $d = 3$, $n = 12$, $m = 24$	9751.6	3060.1	3.2
Matrix minors, $n = 5$, 7×7, $r = 3$	52.6	66.6	0.8
Matrix minors, $n = 6$, 7×7, $r = 3$	556.5	779.1	0.7
Matrix minors, $n = 6$, 8×8, $r = 3$	1257.9	1714.0	0.7
Matrix minors, $n = 4$, 6×6, $r = 5$	262.7	328.1	0.8
Matrix minors, $n = 4$, 7×7, $r = 6$	2872.2	4299.8	0.7
Matrix minors, $n = 4$, 7×10, $r = 7$	4728.4	5485.8	0.9

(b) Second step (F_4 for an elimination order)

System	Without weights (s)	With weights (s)	Speed-up
Cyclic invariants, $n = 4$	4.0	0.3	13.3
Cyclic invariants, $n = 5$, $k = 12$	2705.8	73.4	36.9
Cyclic invariants, $n = 5$	$>$ 90 000 [b]	370.0	>243
Cyclic invariants, $n = 6$, $k = 14$	35 922.4	2256.2	15.9
Cyclic invariants, $n = 6$, $k = 15$	>300 000 [b]	7426.7	>40
Dihedral invariants, $n = 5$	$>$ 40 000 [b]	18.5	>2162
Generic monomials, $d = 2$, $n = 24$, $m = 48$	216.5	110.9	2.0
Generic monomials, $d = 2$, $n = 25$, $m = 50$	31 135.2	16 352.2	1.9
Generic monomials, $d = 2$, $n = 26$, $m = 52$	14 919.2	8142.8	1.8
Generic monomials, $d = 2$, $n = 27$, $m = 54$	5645.8	4619.0	1.2
Generic monomials, $d = 3$, $n = 11$, $m = 22$	370.1	193.1	1.9
Generic monomials, $d = 3$, $n = 12$, $m = 24$	4527.2	2904.6	1.6
Matrix minors, $n = 7$, 7×7, $r = 3$	41 220.0	26 340.0	1.6
Matrix minors, $n = 7$, 8×8, $r = 3$	48 000.0	18 060.0	2.7
Matrix minors, $n = 8$, 8×8, $r = 3$	711 690.0	390 235.0	1.8
Matrix minors, $n = 4$, 6×6, $r = 5$	613.9	325.4	1.9
Matrix minors, $n = 4$, 7×7, $r = 6$	8059.4	3955.5	2.0
Matrix minors, $n = 4$, 7×10, $r = 7$	71 067.8	32 721.5	2.2

b. Memory usage was over 120 GB.

(c) Direct strategy

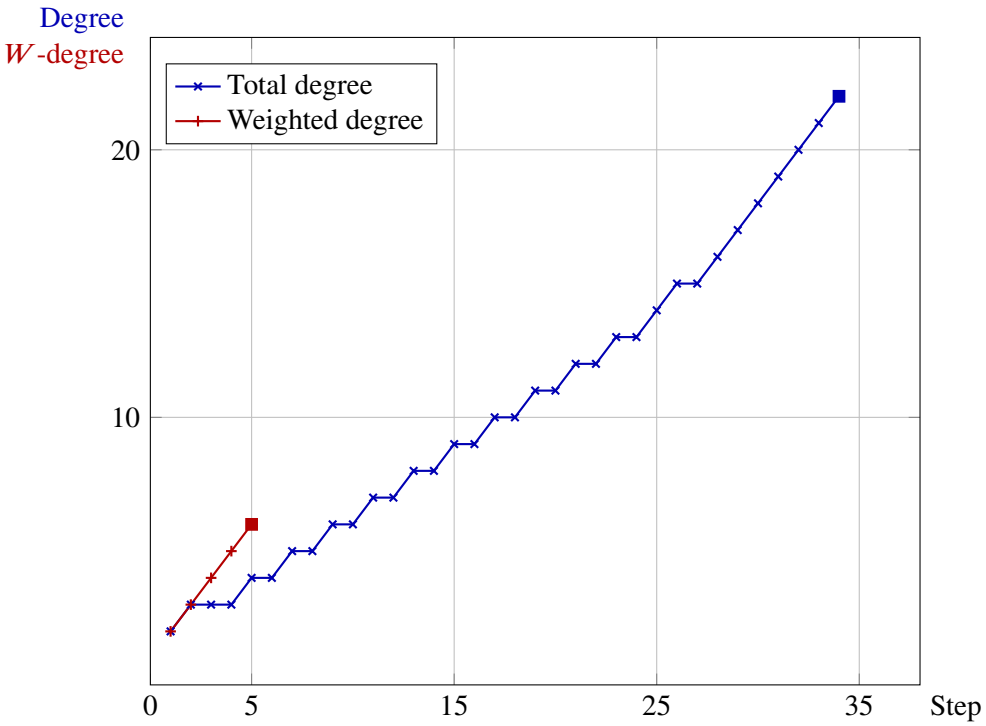

Figure 3.9: Degree at each reduction step in a run of Algorithm F_4 on a polynomial inversion systems (relations between 50 monomials of degree 2 in 25 variables)

Once again, using the weighted degree instead of the total degree makes the system regular in the affine sense, and there is no degree falls with the weighted degree. In this case, this newly found regularity also makes the final degree reached much smaller, which justifies the ridiculous speed-ups that we observed on this class of systems (second block in Table 3.6a).

Chapter 4

Real root classification for determinants Application to contrast optimization

The results presented in this chapter are extracted from a joint work with Bernard Bonnard, Jean-Charles Faugère, Alain Jacquemard and Mohab Safey El Din. They have been published in [Bon+16].

In this chapter, we consider the algorithmic problem of classifying the singularities of real parameterized determinantal systems. This problem has direct applications, for example in contrast optimization for Nuclear Magnetic Resonance imagery. In its most general setting, this example is out of reach of existing implementations of algorithms for real roots classification problems. We describe a computational strategy refining these classic tools in order to take advantage of the structure of determinantal systems. The new algorithm is able to find the classification for the general case of the contrast optimization problem.

4.1. Introduction

Real roots classification and determinantal systems In this chapter, we consider a real roots classification problem for determinantal ideals. A real roots classification problem can be stated as follows: let V be an algebraic variety in $\mathbb{C}^n \times \mathbb{C}^t$, where the n first coordinates are variables and the t last are parameters, and let B be a real semi-algebraic set in $\mathbb{R}^n \times \mathbb{R}^t$. Let $\pi : \mathbb{C}^n \times \mathbb{C}^t \longrightarrow \mathbb{C}^t$. The goal of roots classification is to subdivide $\mathbb{R}^t$ into areas where the cardinality of the fibers of the restriction of π to $V \cap B \subset \mathbb{R}^{n+t}$ is constant. More precisely, we want to find connected open subsets $C_1, \ldots, C_s \subset \mathbb{R}^t$ for the Euclidean topology such that:

1. $C_1 \cup \cdots \cup C_s$ is dense in $\mathbb{R}^t$
2. For any $i \in \{1, \ldots, s\}$, there exists $c_i \in \mathbb{N}$ such that for any $\mathbf{g} \in C_i$, $\#(V \cap B \cap \pi^{-1}(\mathbf{g})) = c_i$.

Obviously, this can only be done if the generic fiber of π on $V \cap B$ is finite.

Here, the variety V shall be a determinantal variety, as defined in Section 1.5. To fix the notations, let M be a $k \times k$ matrix whose coefficients are polynomials in $\mathbb{Q}[X_1, \ldots, X_n, G_1, \ldots, G_t]$. As above, let $\pi : \mathbb{C}^n \times \mathbb{C}^t \longrightarrow \mathbb{C}^t$.

Fix $r \in \{1, \ldots, k-1\}$. Let V_r be the r'th determinantal variety associated with M (that is the set of points of $\mathbb{C}^{n+t}$ at which M has rank at most r), and V the union of the singular locus of V_r and of the critical points of π restricted to V_r.

Figure 4.1: Contrast optimization: both pictures were taken using the same setup of two tubes, the inner one containing deoxygenated blood and the outer one containing oxygenated blood. The left-hand picture is the picture obtained with no contrast optimization; the right-hand picture is contrast-optimized by saturating the inner sample.

Recall from Proposition 1.92 that generically, this variety is V_{r-1}, and it is equidimensional with dimension $n + t - (k - r + 1)^2$. Further assume that

$$n = (k - r + 1)^2,$$

so that generically, V is t-equidimensional. In this case, generically, there exists a Zariski-open set $O \subset \mathbb{C}^t$ such that $V \cap \pi^{-1}(\mathbf{g})$ is a non-empty finite set for any $\mathbf{g} \in O$.

As before, let B be a real semi-algebraic set in $\mathbb{R}^n \times \mathbb{R}^t$, and assume that:

1. B has non-empty interior (for the euclidean topology)
2. the projection π restricted to $V \cap B$ is proper.

We want to design an algorithm for the roots classification problem specialized to $V \cap B$.

Application to contrast optimization This problem arose from a control theory study of the dynamics of contrast optimization in Nuclear Magnetic Resonance (MNR) imaging.

Nuclear Magnetic Resonance (NMR) is a powerful tool in medical imagery. In order to distinguish two biological matters on a picture, it is required to optimize the contrast between the two matters. Because of its importance in medical sciences, this contrast imaging problem has received a lot of attention. The pioneering work of [Bon+12] has established geometric optimal control techniques as a major tool for designing optimal control strategies for the problem of improving the contrast.

These strategies depend on the biological matters under study. In NMR imagery the main physical parameters involved are the longitudinal and transversal relaxation times of each matter. This approach is formalized in [Bon+13], and it leads to a real root classification problem for a determinantal system, where the size of the matrix is $k = 4$ and the number of variables is $n = 4$.

The number of parameters is initially 4: the longitudinal and the transversal relaxation times, for 2 biological matters. Because the system is homogeneous and physical constraints ensure that the parameters are nonzero, we may set one of them to 1, reducing the problem to $t = 3$. An

important particular case is when one of the matters is water; then the corresponding relaxation times are equal, and the number of parameters is $t = 2$.

We also mention that the optimal control problems in [Chy+15] lead to algebraic classification problems with similar structures.

State of the art The modeling through an optimal control problem is introduced in [Bon+13]. The so-called Bloch modeling and saturation method for tackling this problem is developed therein.

In [Bon+13], four experimental important cases are studied (all parameters of the classification problem are fixed). Among other properties, it has been observed there that the number of singularities is constant when water is involved. This led to the following questions:

1. Is this number of singularities preserved for any choice of second matter?
2. If not, how many different classes of pairs of matters can we distinguish through the analysis of those singularities?

Answering these questions leads to the real root classification problem described above. Symbolic computation techniques are good candidates to solve them.

Properties of Cylindrical Algebraic Decomposition (CAD) adapted to a given polynomial family allow to solve real root classification problems. Hence the CAD algorithm [Col75] can be used in our context. However, the complexity of computing a CAD is doubly exponential in the number of variables ([BD07; DH88]); its implementations are usually limited to non-trivial problems involving 4 variables and cannot tackle our application.

The complexity of computing a CAD can be much improved when taking into account equational constraints (see e.g. [McC99]). In the context of real root classification problems, this leads us to take advantage of the presence of equations to compute closed sets in the parameter space ($\mathbb{R}^t$ using our notation) containing the boundaries of the regions $C_1, \ldots, C_\ell$, hence substituting the recursive (doubly exponential) projection steps of CAD with more involved projection techniques. In the past ten years, several works have focused on this problem [LR07; YHX01] using various computer algebra tools such as Gröbner bases, regular chains, etc. We also mention [RT15] which uses evaluation/interpolation techniques to compute those closed sets in the parameters space.

While the implementation of [YHX01] is able to solve our classification problem for the case of water, none of the implementations were able to classify the singular locus of V in the general case (the number of parameters is 3).

Our strategy is based on those of [LR07] and [YHX01], adapted to the determinantal case: by exploiting properties of sets defined by minors of matrices with polynomial entries, we can obtain a more regular behavior for the algorithms. As such, our strategy is less general than existing ones, but is able to compute the partition of the parameter space in the general case.

Such structures have been used for computing sample points in each connected component of the real trace of determinantal varieties [HNS15b; HNS15c; HNS16] or for solving linear matrix inequalities [HNS15a]. These works are based on dedicated strategies for computing critical loci of some projections restricted to determinantal varieties. Such computations are naturally related to real root classification problems and real quantifier elimination (see e.g. [HS12]).

Finally, the computations rely on Gröbner bases. Several works have focused on designing strategies improving the regularity of Gröbner basis computations on determinantal systems [FSS13], and in particular when computing singular and critical loci of determinantal varieties [FSS12; Spa14]. These strategies involve modelling the problem using incidence varieties instead of minors.

Main results The main results of this chapter are twofold:

- an algorithm solving the special real root classification problem described above and which exploits the determinantal structure of the input data arising in contrast imaging problems;
- its successful use for solving the challenging application to contrast problem in the general case.

We start by describing our algorithmic contribution. Recall that we are given a matrix, denoted by M, with polynomial entries. As in [HNS15a; HNS15b; HNS15c; HNS16], it is based on splitting computations according to the rank of M.

More precisely, in order to solve the real root classification problem, we need to identify where the number of real solutions inside B of the determinantal system describing V changes depending on the values of the parameters. In particular, this problem involves inequalities defining a semi-algebraic set with non-empty interior. In this context, we use standard tools from real geometry, such as Thom's first isotopy lemma, which reduce our classification problem to computing the singular points of V, the critical points of the projection of the parameter space restricted to V, and the intersection of V with the boundary of the semi-algebraic set B.

This computation may be difficult because generically, the variety V_r has singularities corresponding to points where $\operatorname{rank}(M) < r$ (Proposition 1.92). Hence, observe that the variety V is naturally split according to the rank of M. This is the very basic idea on which our algorithm relies: we compute critical loci of the projection on the parameter space restricted to the variety V by distinguishing those points at which M has rank less than r from those at which M has rank exactly r.

Our algorithms need to compute projection of algebraic sets, which is done using elimination algorithms such as Gröbner bases or triangular sets for example. We have performed experiments for both these tools, using the package FGb ([Fau10]) in Maple and an implementation of F_5 ([Fau02]) for Gröbner bases, and using the package RegularChains ([LMX05]) in Maple for triangular sets.

Regarding the contrast imaging problem, we illustrate the behaviour of our algorithm in the case of water ($t = 2$), giving the whole classification. Using Gröbner bases to perform the eliminations, the computation takes 10 s on an 2 GHz Intel Xeon CPU. The RealRootClassification command of the Maple RegularChain library needs 1600 s to find this classification.

We also ran our algorithm on the general case ($t = 3$). While none of the available implementations is able to tackle this classification problem directly, ours can find the polynomials separating the open sets C_i within 4 h using FGb, or 2 min using F_5, and the projection step of the CAD can be done in 4 h. We see similar speed-ups when using triangular sets to perform the elimination.

This illustrates how our dedicated algorithms take advantage of the special structure of the problem, to achieve speed-ups when compared with more general techniques.

We give an overview of the results at the end of the chapter[1]. The source code is also given in Appendix A.2.

Structure of the chapter In Section 4.2, we present the mathematical background around NMR imagery and the contrast problem. Section 4.3 deals with the dedicated classification algorithm. Finally, in Section 4.4, we report on experimental results obtained when solving the application to the contrast imaging problem.

4.2. Modeling the dynamics of the contrast optimization

The model we describe below has been introduced in [Bon+12] in order to apply techniques from geometric optimal control theory to the control of the spin dynamics by NMR. Up to some normalization, each spin 1/2 particle is governed by the Bloch equation

$$\begin{cases} \dot{x} = -\Gamma x + u_y z \\ \dot{y} = -\Gamma y - u_x z \\ \dot{z} = \gamma(1 - z) + u_x y - u_y x, \end{cases}$$

where the *state variable* $q = (x, y, z)$ represents the magnetization vector which must lie in the *Bloch ball* defined by $|q| \leq 1$, and the *parameters* (Γ, γ) are related to the physical relaxation times. The parameters must also satisfy $2\Gamma \geq \gamma > 0$. The *control* $u = (u_x, u_y)$ represents the magnetic field whose magnitude is bounded by a maximum value μ.

In the context of the contrast imaging problem, this leads to the simultaneous control of two non-interacting spins with different relaxation time parameters. The *contrast by saturation method* consists in bringing the magnetization vector of the first spin toward the center of the Bloch ball while maximizing the modulus of the magnetization vector of the other matter. The matter with a zero magnetization is black on the picture, while the other matter with a maximum modulus of the magnetization vector is bright.

Using the symmetry of revolution [Bon+13] which allows to eliminate one state variable for each matter, we obtain the system

$$\begin{cases} \dot{y_1} = -\Gamma_1\, y_1 - u_x\, z_1 \\ \dot{z_1} = \gamma_1\,(1 - z_1) + u_x\, y_1 \\ \dot{y_2} = -\Gamma_2\, y_1 - u_x\, z_2 \\ \dot{z_2} = \gamma_2\,(1 - z_2) + u_x\, y_2, \end{cases} \qquad |u| \leq \mu \tag{4.1}$$

and the optimal control problem is: starting from the equilibrium point $N = ((0, 1), (0, 1))$, saturate the first spin, that is $q_1(T) = 0$, where T is the transfer time while maximizing $|q_2(T)|^2$, where $|q_2(T)|$ represents the final contrast. It is a standard Mayer problem in optimal control. It has been studied in [Bon+12] through the analysis of the Hamiltonian dynamics given by the Pontryagin Maximum Principle [Pon+62]. We summarize this analysis below.

[1]The full results, together with the source code which produced them are available at: *mercurey.gforge.inria.fr*

Writing (4.1) as $\dot{q} = F(q) + u\,G(q), \quad |u| \leq \mu$, the optimality conditions associated with the Maximum Principle lead us to construct the optimal solution as a concatenation of *bang-arcs* and *singular arcs*. Bang-arcs are defined by setting the control to $u = \pm\mu$, and singular arcs are solutions u_s of $X_e = F + u_s G$. In the latter case, the control u_s is the rational fraction $-D'/D$ with

$$\begin{aligned} D &= \det(F, G, [G, F], [[G, F], G]) \\ D' &= \det(F, G, [G, F], [[G, F], F]), \end{aligned}$$

where $[\ ,\ ]$ denotes the Lie bracket of vector fields. Explicitly, with $d_i = \gamma_i - \Gamma_i$ ($i \in \{1, 2\}$),

$$D = \det \begin{bmatrix} -\Gamma_1 y_1 & -z_1 - 1 & d_1 z_1 - \Gamma_1 & 2\,d_1 y_1 \\ -\gamma_1 z_1 & y_1 & d_1 y_1 & -2\,d_1 z_1 + \Gamma_1 - d_1 \\ -\Gamma_2 y_2 & -z_2 - 1 & d_2 z_2 - \Gamma_2 & 2\,d_2 y_2 \\ -\gamma_2 z_2 & y_2 & d_2 y_2 & -2\,d_2 z_2 + \Gamma_2 - d_2 \end{bmatrix}.$$

The localization of the singularities of $\{D = 0\}$ inside the Bloch ball is important to understand the geometry of the hypersurface, as well as the dynamics of the vector field X_e which is closely linked to the presence of such singularities.

4.3. Algorithm

4.3.1. Classification strategy

We consider the polynomial algebra $\mathbb{Q}[\mathbf{X}, \mathbf{G}]$ with variables $\mathbf{X} = (X_1, \ldots, X_n)$ and parameters $\mathbf{G} = (G_1, \ldots, G_t)$. Let F and H be families of polynomials in $\mathbb{Q}[\mathbf{X}, \mathbf{G}]$. Let $V_{\mathbb{R}} = V_{\mathbb{R}}(F), V = V_{\mathbb{C}}(F)$ be the set of zeroes of F in $\mathbb{R}^{n+t}$ and in $\mathbb{C}^{n+t}$ respectively. Let B be the closed semi-algebraic set defined by H:

$$B = \{(\mathbf{x}, \mathbf{g}) \in \mathbb{R}^{n+t} \mid \forall h \in H, h(\mathbf{x}, \mathbf{g}) \leq 0\},$$

and let $B_0 = \bigcup_{h \in H} V_{\mathbb{C}}(h)$. Let $\pi : \mathbb{C}^{n+t} \longrightarrow \mathbb{C}^t$ be the projection onto the affine space with coordinates $\mathbf{G}$. Let $\mathrm{sing}(V)$ be the singular locus of V, $\mathrm{crit}(\pi, V)$ be the set of critical points of π restricted to V, and $K(\pi, V) = \pi(\mathrm{sing}(V) \cup \mathrm{crit}(\pi, V)) \cap \mathbb{R}^t$.

Given a subset A of a real or complex affine space, $\overline{A}$ and ∂A are used to denote the closure and the boundary of A for the Euclidean topology respectively.

Assume that the following hypotheses are satisfied:

$\mathcal{H}1$. There exists a nonempty Zariski-open subset $\mathcal{O}_1$ of $\mathbb{C}^t$ such that for all $\mathbf{g} \in \mathcal{O}_1$, the fiber $V \cap \pi^{-1}(\mathbf{g})$ is a nonempty finite subset of $\mathbb{C}^{n+t}$;

$\mathcal{H}2$. The restriction of the projection π to B is proper (Definition 1.81);

$\mathcal{H}3$. The intersection $V \cap B_0$ has dimension at most $t - 1$ in $\mathbb{C}^{n+t}$;

$\mathcal{H}4$. The variety V is equidimensional of dimension t.

We want to find a non-zero polynomial $P \in \mathbb{Q}[\mathbf{G}]$ such that, on each connected component U of $\mathbb{R}^t \setminus V_{\mathbb{R}}(P)$, for $\mathbf{g} \in U$, the cardinality of $V \cap B \cap \pi^{-1}(\mathbf{g})$ does not depend on $\mathbf{g}$.

In Lemma 4.1, we describe a well-known strategy for computing these objects (see for example [LR07; YHX01]).

Lemma 4.1. *Let F and H be polynomial systems satisfying hypotheses $\mathcal{H}1$, $\mathcal{H}2$, $\mathcal{H}3$ and $\mathcal{H}4$. Let $C_B = \pi(V \cap B_0)$, U a non-empty connected open subset of $\mathbb{R}^t$ which does not meet $C_B \cup K(\pi, V)$, and $\mathbf{g} \in U$. Then $V \cap \pi^{-1}(\mathbf{g})$ is finite, and for any $\mathbf{g}' \in U$, $\#\left(V \cap \pi^{-1}(\mathbf{g}')\right) = \#\left(V \cap \pi^{-1}(\mathbf{g})\right)$.*

Proof. We will construct a Whitney stratification of $V \cap B$ ([BCR98, Def. 9.7.1]) with certain properties. First note that since V is t-equidimensional by hypothesis $\mathcal{H}4$, V has real dimension at most t (Proposition 1.60). Let $\mathcal{S}_{=t}$ be the intersection of the points where $V_{\mathbb{R}}$ has local dimension t and of the interior of B. There exists a Whitney stratification $(\mathcal{S}_i)$ of the semi-algebraic set $V \cap B$ such that $\mathcal{S}_{=t}$ is the union of strata of dimension t ([BCR98, Th. 9.7.11]). Let $\mathcal{S}_{<t}$ be the union of the other strata, they all have real dimension less than t. By construction, this is a semi-algebraic set which is the union of $(V \cap \partial B) \subset (V \cap B_0)$ and of the singular locus of $V \cap B$, and it has dimension less than t (using hypothesis $\mathcal{H}3$ for $V \cap B_0$). Its image through π has dimension less than t, and so it has codimension at least 1.

Now consider $\mathcal{S}_{=t}$. By hypothesis $\mathcal{H}1$, for any $\mathbf{g} \in \mathcal{O}_1$, $\pi^{-1}(g) \cap V$ is non-empty, hence $\pi(V)$ contains the non-empty Zariski-open set $\mathcal{O}_1 \subset \mathbb{C}^t$. The intersection $\mathcal{O}_1 \cap \mathbb{R}^t$ is a non-empty Zariski-open set of $\mathbb{R}^t$, contained in $\pi(V)$, hence $\pi(V) \cap \mathbb{R}^t$ has real dimension t. Let U_0 be its interior. The subset $\mathcal{S}_{=t} \cap \pi^{-1}(U_0)$ is a locally closed semi-algebraic set. If it is empty, then there is nothing to prove. Otherwise, by construction it has dimension t; and the projection π restricted to this subspace is proper, by hypothesis $\mathcal{H}2$. Thom's isotopy lemma (Theorem 1.82) states that for any nonempty connected open set U of $\mathbb{R}^t$ not meeting $K(\pi, V)$, and for any $\mathbf{g} \in U$, there exists a semi-algebraic diffeomorphism

$$h = (h_0, \pi) : V \cap B \cap \pi^{-1}(U) \xrightarrow{\sim} \pi^{-1}(\mathbf{g}) \times U.$$

By hypothesis $\mathcal{H}1$, if U is nonempty, $\pi^{-1}(\mathbf{g})$ is finite, and the cardinality of the fibers is constant on U. □

So in order to compute the wanted decomposition of the parameter space, it suffices to compute a polynomial $P \in \mathbb{Q}[\mathbf{G}]$ such that $V(P)$ covers $\pi(V \cap B_0)$ and $K(\pi, V)$.

4.3.2. The determinantal problem

Let k be an integer greater than 1, $r_0 \in \{1, \ldots, k-1\}$, and

$$n = (k - r_0 + 1)^2. \tag{4.2}$$

Let $t \in \mathbb{N}$, and let $M(\mathbf{X}, \mathbf{G})$ be a $k \times k$ matrix with polynomial entries in n variables $\mathbf{X} = (X_1, \ldots, X_n)$ and t parameters $\mathbf{G} = (G_1, \ldots, G_t)$. As before, let $\pi : \mathbb{C}^{n+t} \longrightarrow \mathbb{C}^t$ be the projection onto the affine space with coordinates $\mathbf{G}$.

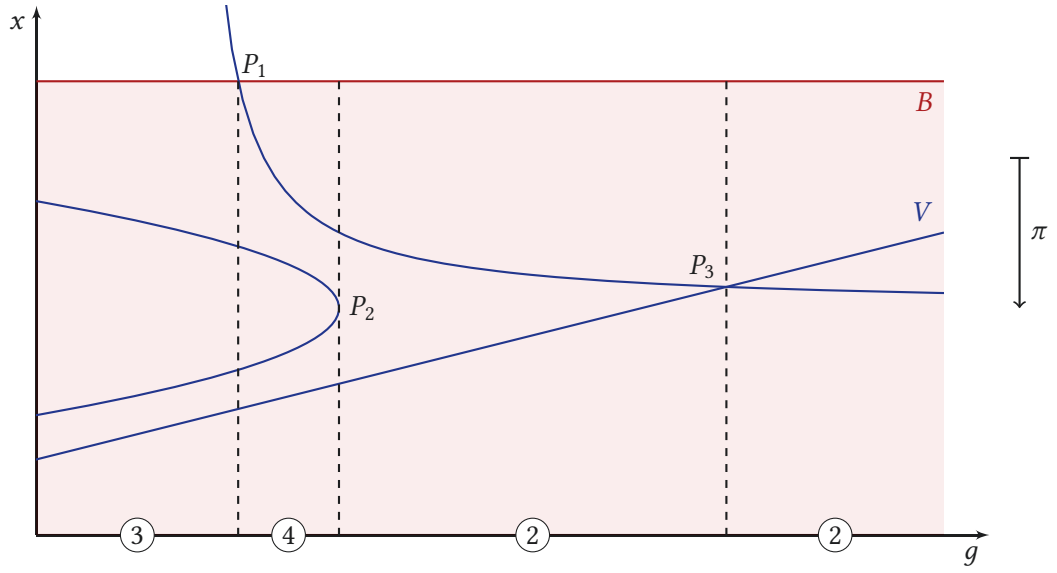

Figure 4.2: Roots classification using Thom's isotopy lemma: we classify the points of the g axis according to the cardinality of the fibers of π restricted to $V \cap B$. The points where this cardinality may change are the projections of P_1, P_2 and P_3: P_1 is in the intersection of V and the boundary of B, P_2 is a critical point of π restricted to V and P_3 is a singular point of V.

Let $\{h(\mathbf{X}, \mathbf{G}) \leq 0 \mid h \in H\}$ be a system of inequalities, with $H \subset \mathbb{Q}[\mathbf{X}, \mathbf{G}]$. We will consider determinantal varieties associated with M, as defined in Section 1.5: for any $r \in \{0, \ldots, k\}$, we define the variety

$$V_r = \{(\mathbf{x}, \mathbf{g}) \in \mathbb{C}^{n+t} \mid \operatorname{rank}(M(\mathbf{x}, \mathbf{g})) \leq r\},$$

and let $V_{-1} = \varnothing$ by convention. Furthermore, we define the constructible set $V_{=r} = V_r \setminus V_{r-1}$, that is the set of points at which the matrix M has rank exactly r.

Let V be the union of the singular locus of V_{r_0} and of the set of critical points of π restricted to V_{r_0}. In other words,

$$\pi(V) = K(\pi, V_{r_0}).$$

We want to classify the cardinality of the real fibers by π of the semi-algebraic set

$$V \cap \{(\mathbf{x}, \mathbf{g}) \mid \forall h \in H, h(\mathbf{x}, \mathbf{g}) \leq 0\}.$$

Assume that V and H satisfy hypotheses $\mathcal{H}1$, $\mathcal{H}2$, $\mathcal{H}3$ and $\mathcal{H}4$. Further assume that:

$\mathcal{H}5$. There exists a non-empty Zariski-open subset $O_2 \subset \mathbb{C}^t$ such that

$$V \cap \pi^{-1}(O_2) = V_{r_0-1} \cap \pi^{-1}(O_2)$$

$\mathcal{H}6$. For any $r \in \{0, \dots, k-1\}$, the ideal defined by the $(r+1)$-minors of M is radical;

$\mathcal{H}7$. For any $r \in \{0, \dots, k-1\}$, the variety V_r is equidimensional with dimension $n + t - (k - r + 1)^2$.

By Proposition 1.92, these properties are generic. Furthermore, hypothesis $\mathcal{H}7$ implies hypothesis $\mathcal{H}4$ (by definition (4.2) of n), and generically it implies hypotheses $\mathcal{H}1$.

In following subsections, we will describe two algorithms DeterminantCritVals and DeterminantBoundary, which, given such a matrix M, a target rank r_0 and inequalities H, compute respectively a polynomial whose zeroes cover $K(\pi, V)$, and a polynomial whose zeroes cover $\pi(V \cap B_0)$. By Lemma 4.1, the zeroes of the product of these polynomials will subdivide the parameter space into connected components where the cardinality of real fibers is constant. These algorithms are probabilistic, because they will rely on the choice of generic linear forms to ensure linear independence. However, the algorithms could be made deterministic by testing that these linear forms are generic enough for our purpose, and repeating the random draw otherwise.

The algorithms will also need to compute the projection of algebraic sets onto coordinate subspaces. For this purpose, we assume that we are given a routine Elimination, which, given a system of polynomials $F \subset \mathbb{Q}[V_1, \dots, V_N]$ and a set of variables $\mathbf{V}' \subset \{V_1, \dots, V_N\}$, computes a system of generators of $\langle F \rangle \cap \mathbb{Q}[\mathbf{V}']$. Such a routine can be implemented using Gröbner bases or regular chains, for example.

4.3.3. Incidence varieties

We decompose the problem depending on the rank of the matrix. For that purpose, we use incidence varieties, as defined in Section 1.5.3.

Recall that the incidence variety of rank r associated with M is the variety $\mathcal{V}_r \subset \mathbb{C}^{n+t} \times (\mathbb{P}^{k-1}(\mathbb{C}))^{k-r}$ defined by:

$$M \cdot \begin{bmatrix} y_{1,1} & \dots & y_{1,k-r} \\ \vdots & & \vdots \\ y_{k,1} & \dots & y_{k,k-r} \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} \tag{4.3}$$

with the additional condition that the matrix $(y_{i,j})$ has rank $k - r$.

Let $(u_{1,1}, \dots, u_{k-r,k}) \in \mathbb{C}^{k(k-r)}$, we define the variety $\mathcal{V}'_{r,\mathbf{u}}$ as the intersection of $\mathcal{V}_r$ and the complex solutions of

$$\begin{bmatrix} u_{1,1} & \dots & u_{1,k} \\ \vdots & & \vdots \\ u_{k-r,1} & \dots & u_{k-r,k} \end{bmatrix} \cdot \begin{bmatrix} y_{1,1} & \dots & y_{1,k-r} \\ \vdots & & \vdots \\ y_{k,1} & \dots & y_{k,k-r} \end{bmatrix} = \mathrm{Id}_{k-r}, \tag{4.4}$$

this encodes that the matrix $(y_{i,j})$ has rank $k-r$ and it fixes an affine chart for the homogeneous coordinates of the columns of $(y_{i,j})$.

In the rest of Section 4.3, $\mathcal{F}_{r,\mathbf{u}}$ denotes the union of Equations (4.3) and (4.4).

Proposition 4.2. *Let $r_1 = r_0 - 1$. Let $\varphi : \mathcal{V}'_{r_1,\mathbf{u}} \longrightarrow \mathbb{C}^t$ be the projection onto the affine space with coordinates $\mathbf{G}$. Assuming that hypotheses $\mathcal{H}1$ to $\mathcal{H}7$ hold, there exists a Zariski-open subset $\mathcal{U} \subset \mathbb{C}^{k(k-r_1)}$ such that if $\mathbf{u} \in \mathcal{U} \cap \mathbb{Q}^{k(k-r_1)}$, $\overline{K(\pi, V_{r_1})} = \overline{K(\varphi, \mathcal{V}'_{r_1,\mathbf{u}})}$.*

Proof. Let $P = (\mathbf{x}, \mathbf{g}, \mathbf{y}) \in \mathcal{V}'_{r_1,\mathbf{u}}$.

If $M(\mathbf{x}, \mathbf{g})$ has rank less than r_1, then by Proposition 1.91, $(\mathbf{x}, \mathbf{g}) \in \mathrm{sing}(V_{r_1})$, hence $\mathbf{g} \in K(\pi, V_{r_1})$.

Since $M(\mathbf{x}, \mathbf{g})$ has rank less than r_1, its kernel L_1 has dimension at least $k - r_1 + 1$. Equations (4.4) encode that the vectors $\mathbf{y}_i$ given by the columns of matrix $(y_{i,j})$ generate a r_1-dimensional vector space L_2. So there exists $\mathbf{y}_0 \in L_1 \cap L_2$, and for all $a \in \mathbb{C}$, $(\mathbf{x}, \mathbf{g}, \mathbf{y}_1 + a\mathbf{y}_0, \mathbf{y}_2)$ belongs to the fiber above $(\mathbf{x}, \mathbf{g})$ in $\mathcal{V}_{r_1}$. So this fiber has dimension at least 1, while the generic fiber has dimension 0 by hypothesis $\mathcal{H}1$. So $(\mathbf{x}, \mathbf{g})$ is a critical value of the projection of $\mathcal{V}_{r_1}$ onto $\mathbb{R}^{n+t}$, hence $(\mathbf{g}) \in K(\varphi, \mathcal{V}'_{r_1,\mathbf{u}})$.

So we may assume that $M(\mathbf{x}, \mathbf{g})$ has rank exactly r_1. There is a $r_1 \times r_1$ submatrix A of $M(\mathbf{x}, \mathbf{g})$ which is invertible, without loss of generality we may assume that it is the top-left $r_1 \times r_1$ submatrix. In an open neighborhood of $(\mathbf{x}, \mathbf{g})$, $V_{=r_1}$ is described by the vanishing of the entries of M/A, that is the determinants of the $(r_1 + 1) \times (r_1 + 1)$ submatrices containing A. The same computations as in the proof of Lemma 1.100 give the following equations describing $\mathcal{V}'_{r_1,\mathbf{u}}$ in the open neighborhood of $(\mathbf{x}, \mathbf{g}, \mathbf{y})$ where $\Delta = \det(A)$ does not vanish:

$$\begin{cases} M/A = 0 \\ Y_{(2)} = (U_{(2)} - U_{(1)}A^{-1}B)^{-1} \\ Y_{(1)} = \frac{-1}{\Delta}A^{-1}BY_{(2)} \end{cases} \tag{4.5}$$

and the truncated Jacobian matrix in $(\mathbf{X}, \mathbf{Y})$ of this system can be written

$$\begin{bmatrix} \mathrm{Jac}_{\mathbf{X}}(M/A) & 0 & 0 \\ \star & \mathrm{Id}_{(k-r_1)^2} & \star \\ \star & 0 & \mathrm{Id}_{r_1(k-r_1)} \end{bmatrix}$$

where $\mathrm{Jac}_{\mathbf{X}}(M/A)$ is the truncated Jacobian matrix in $\mathbf{X}$ of the $(k - r_1)^2$ entries of M/A, which define $V_{r_1} \setminus \{(\mathbf{x}, \mathbf{g} \mid \Delta = 0)\}$ in $\mathbb{C}^{n+t}$. By hypothesis $\mathcal{H}6$, the ideal defined by the entries of M/A, which is a subideal of the ideal of all $(r_1 + 1)$-minors of M, is radical. Since the Schur complement appears by multiplication with invertible matrices with entries in the localized ring $\mathbb{Q}[\mathbf{X}, \mathbf{g}]_\Delta$ (using the same notations as in the proof of Lemma 1.100):

$$\begin{bmatrix} \mathrm{Id}_{r_1} & 0 \\ -C & \mathrm{Id}_{k-r_1} \end{bmatrix} \cdot \begin{bmatrix} A^{-1} & 0 \\ 0 & \mathrm{Id}_{k-r_1} \end{bmatrix} \cdot M = \begin{bmatrix} \mathrm{Id}_{r_1} & A^{-1}B \\ 0 & M/A \end{bmatrix},$$

Equations (4.5) describe the localization of $\langle \mathcal{F}_{r_1,\mathbf{u}} \rangle$ in $\mathbb{Q}[\mathbf{X}, \mathbf{g}]_\Delta$, so this ideal is radical as well. So we can use the Jacobian criterion on V_{r_1} near $(\mathbf{x}, \mathbf{g})$ and on $\mathcal{V}'_{r_1,\mathbf{u}}$ near $(\mathbf{x}, \mathbf{g}, \mathbf{y})$. Both Jacobians matrices have the same rank and both varieties have the same local codimension $(k - r_1)^2$ (by Lemma 1.100 and hypothesis $\mathcal{H}7$), so

$$K(\pi, V_{=r_1}) \cap \varphi(\mathcal{V}'_{r_1,\mathbf{u}}) = K(\varphi, \mathcal{V}'_{r_1,\mathbf{u}}) \cap \pi(V_{=r_1})$$

The image $\varphi(\mathcal{V}'_{r_1,\mathbf{u}}) \cap \pi(V_{=r_1})$ is a Zariski-open subset $\mathcal{O}_{\mathbf{u}}$ of $\pi(V_{=r_1})$. It remains to prove that if $\mathbf{u}$ is sufficiently generic, then all irreducible components of $K(\pi, V_{=r_1})$ meet this open subset.

Let $C_1, \dots, C_a$ be these irreducible components, and let

$$(\mathbf{x}_1, \mathbf{g}_1) \in \pi^{-1}(C_1), \dots, (\mathbf{x}_a, \mathbf{g}_a) \in \pi^{-1}(C_a).$$

For any $(\mathbf{x}, \mathbf{g}) \in V_{=r_1}$, the proof of [HNS15b, Prop. 4, Sec. 6] shows that there exists a non-empty Zariski-open subset $\mathcal{U}_{(\mathbf{x},\mathbf{g})} \subset \mathbb{C}^{k(k-r_1)}$ such that if $\mathbf{u} \in \mathcal{U}_{(\mathbf{x},\mathbf{g})} \cap \mathbb{Q}^{k(k-r_1)}$, then $(\mathbf{x}, \mathbf{g}) \in O_{\mathbf{u}}$; namely, $\mathcal{U}_{(\mathbf{x},\mathbf{g})}$ is the set of $\mathbf{u}$ such that

$$\operatorname{rank} \begin{bmatrix} M(\mathbf{x}, \mathbf{g}) \\ (u_{i,j}) \end{bmatrix} = k.$$

Taking the finite intersection of the non-empty Zariski-open subsets $\mathcal{U}_{(\mathbf{x}_i,\mathbf{g}_i)}$ for $i \in \{1, \dots, a\}$ yields the wanted subset $\mathcal{U}$. □

4.3.4. Locus of rank exactly r_0

Recall that by $\mathcal{H}5$, $\pi(V \cap V_{=r_0})$ has codimension at least 1, and that we want to compute a polynomial whose zeroes cover $K(\pi, V)$ and $\pi(V \cap B_0)$. So we may multiply the result by the equation of one hypersurface covering $\pi(V \cap V_{=r_0})$, it will naturally cover $\pi(V \cap V_{=r_0}) \cap K(\pi, V)$ and $\pi(V \cap V_{=r_0}) \cap \pi(V \cap B_0)$.

Algorithm 4.1 RankExactly

Input: $M \subset \mathbb{Q}[\mathbf{X}, \mathbf{G}]^{k \times k}$, $r_0 \in \{1, \dots, k-1\}$
Output: $P_1 \in \mathbb{Q}[\mathbf{G}] \setminus \{0\}$ *s.t.* $\pi(V \cap V_{=r_0}) \subset V(P_1)$

```
res ← 1
F_{V,0} ← {(r_0 + 1)-minors of M}
J ← Jac_X(F_{V,0})
F_{V,1} ← F_{V,0} ∪ {(k − r_0)^2-minors of J}
Pick at random u_1, ..., u_{k(k−r_0)} = u ∈ Q^{k(k−r_0)}
F_0 ← F_{k−r_0,u}
{M_1, ..., M_N} ← {r_0-minors of M}
for i in {1, ..., N} do
    F_1 ← F_0 ∪ F_{V,1} ∪ {M_1, ..., M_{i−1}, u · M_i − 1}
    G ← Elimination(F_1, {u, X, Y})
    Multiply res by 1 polynomial from G
end for
return res
```

4.3.5. Singularities

Algorithm 4.2 DeterminantCritVals

Input: $M \subset \mathbb{Q}[\mathbf{X}, \mathbf{G}]^{k\times k}$, $r_0 \in \{1, \dots, k-1\}$
Output: $P_c \in \mathbb{Q}[\mathbf{G}] \setminus \{0\}$ *s.t.* $K(\pi, V) \subset V(P_c)$
1: res $\leftarrow$ RankExactly(M, r_0)
2: Pick at random $u_1, \dots, u_{k(k-r_0+1)} = \mathbf{u} \in \mathbb{Q}^{k(k-r_0+1)}$
3: $F_0 \leftarrow \mathcal{F}_{r_0-1,\mathbf{u}}$
4: $J \leftarrow \mathrm{Jac}_{\mathbf{X},\mathbf{Y}}(F_0)$
5: $F_1 \leftarrow F_0 \cup \{k(k-r_0+1)+(k-r_0+1)^2\text{-minors of } J\}$
6: $G \leftarrow$ Elimination$(F_1, \{\mathbf{X}, \mathbf{Y}\})$
7: Multiply res by 1 polynomial from G
8: **return** res

Proposition 4.3. *Algorithm DeterminantCritVals is correct.*

Proof. By definition, $V \subset V_{r_0}$. Using the decomposition $V_{r_0} = V_{=r_0} \cup V_{r_0-1}$, we decompose the variety V as $V = \big(V \cap V_{=r_0}\big) \cup \big(V \cap V_{r_0-1}\big)$.

The subspace $\pi(V \cap V_{=r_0})$ is covered by the output of RankExactly, so we may restrict to $V \cap V_{r_0-1}$, which is the whole variety V_{r_0-1} by Lemma 1.91.

By Proposition 4.2, in order to compute $K(\pi, V_{r_0-1})$, we can compute polynomials whose zeroes cover $K(\varphi, \mathcal{V}'_{r_0-1,\mathbf{u}})$ with $\mathbf{u}$ sufficiently generic instead.

By hypotheses $\mathcal{H}6$, $\mathcal{H}7$ and the proof of Proposition 4.2, $\mathcal{V}'_{r_0-1,\mathbf{u}}$ is t-equidimensional and $\mathcal{F}_{r_0-1,\mathbf{u}}$ is a set of generators of its ideal, so we can use the Jacobian criterion to compute equations defining $K(\varphi, \mathcal{V}'_{r_0-1,\mathbf{u}})$. □

4.3.6. Boundary

Algorithm 4.3 DeterminantBoundary

Input: $M \subset \mathbb{Q}[\mathbf{X}, \mathbf{G}]^{k\times k}, r_0 \in \{1, \dots, k-1\}, H \subset \mathbb{Q}[\mathbf{X}, \mathbf{G}]$
Output: $P_b \in \mathbb{Q}[\mathbf{G}] \setminus \{0\}$ *s.t.* $\pi(V \cap B_0) \subset V(P_b)$
1: res $\leftarrow$ RankExactly(M, r_0)
2: Pick at random $u_1, \dots, u_{k(k-r_0+1)} = \mathbf{u} \in \mathbb{Q}^{k(k-r_0+1)}$
3: $F_0 \leftarrow \mathcal{F}_{r_0-1,\mathbf{u}}$
4: **for** h in H **do**
5: $\quad F_1 \leftarrow F_0 \cup \{h\}$
6: $\quad G \leftarrow$ Elimination$(F, \{\mathbf{X}, \mathbf{Y}\})$
7: $\quad$ Multiply res by 1 polynomial from G
8: **end for**
9: **return** res

Proposition 4.4. *Algorithm DeterminantBoundary is correct.*

Proof. As in Section 4.3.5, we write: $V = \big(V \cap V_{=r_0}\big) \cup \big(V \cap V_{r_0-1}\big)$. Since $B_0 = \bigcup_{h\in H} V(h)$, the intersection $V \cap \partial B$ is contained in the union of the varieties $V(\langle F\rangle + \langle h\rangle)$ for h ranging over H, and the equation of the projections can be obtained with polynomial elimination. □

Remark 4.5. For the real root classification problem, the subdivision is given by the product of the outputs of DeterminantCritVals and DeterminantBoundary. In order to avoid repeating computations, we may skip the call to RankExactly in either subroutine (but not both), and initialize res to 1 instead.

4.4. The contrast problem

4.4.1. The case of water

With the notations of Section 4.2, the variety V is the complex algebraic variety defined by

$$D = \frac{\partial D}{\partial y_1} = \frac{\partial D}{\partial y_2} = \frac{\partial D}{\partial z_1} = \frac{\partial D}{\partial z_2} = 0.$$

With the notations of Section 4.3, we want to classify the singularities of the set of points where M has rank at most $r_0 = 3$. Our semi-algebraic constraints are that the solutions are within the Bloch ball, that is

$$\mathcal{B} : \begin{cases} h_1 = y_1^2 + (z_1+1)^2 \le 1 \\ h_2 = y_2^2 + (z_2+1)^2 \le 1. \end{cases}$$

Since the equations are homogeneous in $\Gamma_1, \Gamma_2, \gamma_1, \gamma_2$, and the parameters are supposed to be non-zero, we may normalize by setting $\gamma_1 = 1$. In the case where the first matter is water, we further simplify by setting $\Gamma_1 = \gamma_1 = 1$, leaving free the two parameters Γ_2, γ_2 corresponding to the second matter. We recall that we also assume that $2\,\Gamma_2 \ge \gamma_2$ and that $(\gamma_2, \Gamma_2) \ne (1,1) = (\gamma_1, \Gamma_1)$ (that is, the second matter is not water).

This system satisfies hypotheses $\mathcal{H}1$ to $\mathcal{H}7$.

Theorem 4.6. *Consider the 9 polynomials:*

$$\begin{aligned}
f_1 &= \Gamma_2 - 1 \\
f_2 &= 3\,\Gamma_2 - 2\,\gamma_2 - 1 \\
f_3 &= 3\,\Gamma_2^2 - 5\,\Gamma_2\gamma_2 + \gamma_2^2 + 2\,\Gamma_2 - 2\,\gamma_2 + 1 \\
f_4 &= 2\,\Gamma_2^2 - 5\,\Gamma_2\gamma_2 + 2\,\gamma_2^2 - 2\,\Gamma_2 + 3\,\gamma_2 \\
f_5 &= 2\,\gamma_2^3 - (3\,\Gamma_2 + 11)\,\gamma_2^2 + \big(9\,\Gamma_2 + 6 - 3\,\Gamma_2^2\big)\,\gamma_2 + 2\,\Gamma_2\,(\Gamma_2+2)\,(\Gamma_2 - 1) \\
f_6 &= \Gamma_2 - 2\,\gamma_2 + 1 \\
f_7 &= 2\,\Gamma_2 - \gamma_2 - 1 \\
f_8 &= \gamma_2 - 2 + \Gamma_2 \\
f_9 &= 2\,\Gamma_2^2 - 5\,\Gamma_2\gamma_2 + 2\,\gamma_2^2 + 1
\end{aligned}$$

The zeroes of their product divide the subset of $\mathbb{R}^2$ defined by $2\,\Gamma_2 > \gamma_2 > 0$ into connected components where the cardinality of $V_{\mathbb{R}} \cap \pi^{-1}(\gamma_2, \Gamma_2)$ is constant.

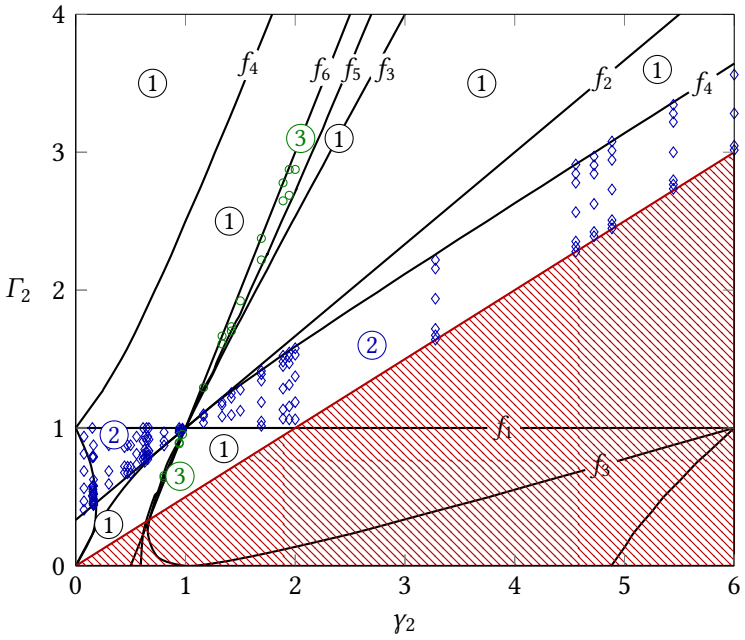

Figure 4.3: Curves involved in the definition of the semi-algebraic set $\mathcal{G}$. Sample points marked with a blue diamond (*resp.* a green circle) are points in $\mathcal{G}_1^- \cup \mathcal{G}_1^+$ (*resp.* $\mathcal{G}_2^- \cup \mathcal{G}_2^+$). Parameters in the red area are physically irrelevant.

Let $\psi : (y_1, z_1, y_2, z_2) \longmapsto (-y_1, z_1, -y_2, z_2)$ be the symmetry fixing $\Pi = \{y_1 = y_2 = 0\}$, and let us consider the semi-algebraic sets (see Figure 4.3):

$$\begin{aligned}
\mathcal{G}_1^- &= \{\gamma_2 < 2\,\Gamma_2, \Gamma_2 < 1, f_2 > 0, f_4 < 0\},\\
\mathcal{G}_1^+ &= \{\gamma_2 < 2\,\Gamma_2, \Gamma_2 > 1, f_2 < 0, f_4 > 0\},\\
\mathcal{G}_2^- &= \{\Gamma_2 < 1, f_6 > 0, f_3 < 0\},\\
\mathcal{G}_2^+ &= \{\Gamma_2 > 1, f_6 < 0, f_5 > 0\},\\
\mathcal{G} &= \mathcal{G}_1^- \cup \mathcal{G}_1^+ \cup \mathcal{G}_2^- \cup \mathcal{G}_2^+.
\end{aligned}$$

Theorem 4.7. *For all (γ_2, Γ_2) such that $2\,\Gamma_2 > \gamma_2 > 0$, the center O of the Bloch ball $\mathcal{B}$ is a singularity of $\{D = 0\}$. And provided $(\gamma_2, \Gamma_2) \in \mathcal{G}$, there exist at most two other singularities:*

1. *provided $(\gamma_2, \Gamma_2) \in \mathcal{G}_1^- \cup \mathcal{G}_1^+$ there is one other singularity lying on $\Pi \cap \mathcal{B}$;*
2. *provided $(\gamma_2, \Gamma_2) \in \mathcal{G}_2^- \cup \mathcal{G}_2^+$ there are two other singularities in $\mathcal{B}$, ψ-symmetric, outside Π.*

The configuration is illustrated in Figures 4.3 and 4.4. Observe that the number of singularities inside $\mathcal{B}$ is an invariant of the contrast problem. Two of the pairs of biological matters studied in [Bon+13], water-cerebrospinal fluid (normalized parameters $[\gamma_2 = \frac{5}{4}, \Gamma_2 = \frac{25}{3}]$) and water-fat (normalized parameters $[\gamma_2 = \frac{25}{2}, \Gamma_2 = 25]$) correspond to points outside $\mathcal{G}$, and their invariant is 1 in both cases (see Figure 4.4). But our results give answers to our guiding questions: there

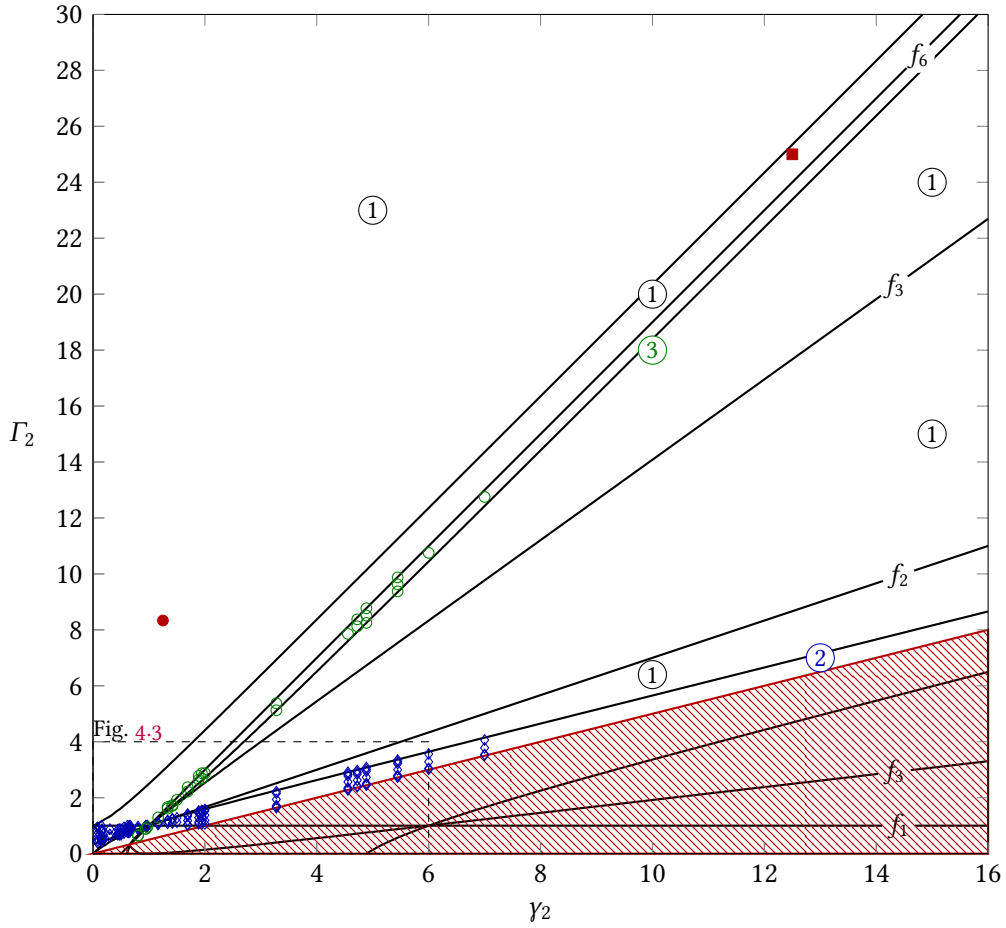

Figure 4.4: Positions of the parameters corresponding to the pairs water-cerebrospinal fluid (red circle) and water-fat (red square) and the set $\mathcal{G}$ (with the same conventions as in Figure 4.3).

exist pairs of matters for which this algebraic invariant can differ; and any pair (water, matter) belongs to one of 3 classes, depending on whether the number of singularities inside $\mathcal{B}$ is 1, 2 or 3.

Proof of Theorem 4.6. Let $V_{=3} = \{p \in \mathbb{C}^4 \times \mathbb{R}^2 \mid \text{rank}(M) = 3\}$ and $V_2 = \{p \in \mathbb{C}^4 \times \mathbb{R}^2 \mid \text{rank}(M) < 3\}$, where $p = (y_1, y_2, z_1, z_2, \gamma_2, \Gamma_2)$. We apply the strategies described in Section 4.3.

We study the generic case $V_2 \cap V$ first. This set does cover a dense subset of $\mathbb{R}^2$. Its intersection with the boundary of $\mathcal{B}$ is given by the vanishing of either h_1 or h_2. The projection on (Γ_2, γ_2) of the set of points of $V_2 \cap V$ such that $h_1 = 0$ is described by $0 = \gamma_2 f_1^2 f_2 f_3$ which gives us polynomials f_1, f_2 and f_3.

The projection on (Γ_2, γ_2) of the set of points of $V_2 \cap V$ such that $h_2 = 0$ is described by $0 = (2\,\Gamma_2 - \gamma_2)\, f_1^2 f_4 f_5$ which gives us new polynomials f_4 and f_5.

System	RegularChain (direct)	Gröbner (direct)	RegularChain (new algo.)	FGb (new algo.)	F5 (new algo.)	CAD
Water	1600 s	100 s		10 s	1 s	50 s
General	>24 h	>24 h	90 × 200 s	46 × 200 s	110 s	4 h (projection step)

Table 4.1: Timings

Next, we consider the incidence variety $\mathcal{V}_2$ associated with the matrix M:

$$M \cdot \begin{bmatrix} \lambda_{1,1} & \lambda_{1,2} \\ \lambda_{2,1} & \lambda_{2,2} \\ \lambda_{3,1} & \lambda_{3,2} \\ \lambda_{4,1} & \lambda_{4,2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

with random linear equations ensuring that the matrix $(\lambda_{i,j})$ has rank 2.

Out of the surface $\gamma_2 = 0$, this affine variety is a complete intersection (it has dimension 2 and it is given by 9 equations in 11 variables, including the saturation by γ_2). The set of critical values of π is described by $0 = (2\,\Gamma_2 - \gamma_2)\,(\Gamma_2 + 1) f_1^2 f_6^2 f_7^2$ which gives us new polynomials f_6 and f_7 ($\Gamma_2 + 1$ has no solutions within our constraint range).

This completes the study of $V \cap V_2$. We now move on to the study of $V \cap V_{=3}$. As described in the algorithm, we define the incidence variety of rank 3 of M, and we saturate successively by the 3-minors of M. Only the first of these subcases is nonempty, and it is described by $0 = (2\,\Gamma_2 - \gamma_2)\, f_8 f_9$ which gives us f_8 and f_9. □

Proof of Theorem 4.7. Observe first by means of a trivial evaluation that O is a singularity of $\{D = 0\}$. We now focus on singularities in $\mathcal{B}^* = \mathcal{B} \setminus \{O\}$. Theorem 4.6 provides a list of 9 polynomials to which we add our constraints $2\,\Gamma_2 \geq \gamma_2 > 0$. Let $\xi = \gamma_2\,\Gamma_2\,(\gamma_2 - 2\,\Gamma_2) \prod_{i=1}^{9} f_i$. The complementary of $\{\xi = 0\}$ is the union of a sequence of connected open semi-algebraic sets where the number of singularities is constant. The routine CylindricalAlgebraicDecompose of the Maple package RegularChains[SemiAlgebraicSetTools] provides 1533 sample points. Excluding those at which ξ vanishes and those outside our physical constraints domain, remains a set K_c of 548 points. At each point of K_c we locate the singularities by computing a Gröbner basis.

We get 165 points of K_c such that there exists at least one singularity in $\mathcal{B}^*$. We have a set K_s of 37 points, each of them corresponding to a couple of ψ-symmetric singularities outside the symmetry plane Π, and a set K_p of 128 points corresponding to a unique singularity on $\Pi \cap \mathcal{B}^*$. For parameters at which ξ does not vanish, the number of singularities in $\mathcal{B}^*$ is at most two.

Points of K_s (*resp.* K_p) are represented in green (*resp.* blue) in Figures 4.3, 4.5 and 4.6. Let us evaluate on K_c the condition ($\Gamma_2 < 1$, $f_2 > 0$, $f_4 < 0$) or ($\Gamma_2 > 1$, $f_2 < 0$, $f_4 > 0$). Indeed the set of points of K_c satisfying this condition coincides with K_p. This proves item 1. The proof of item 2 is similar. □

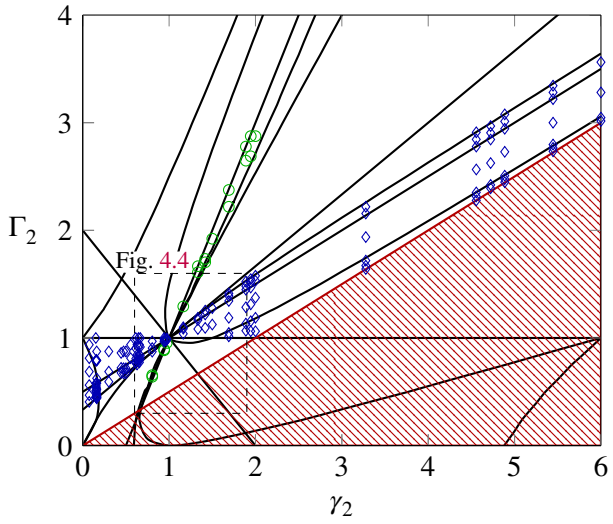

Figure 4.5: The curves involved in the decomposition of the region $\Gamma_2 > 0, \gamma_2 > 0, 2\,\Gamma_2 \geq \gamma_2$ of the parameter space (with the same conventions as in Figure 4.3)

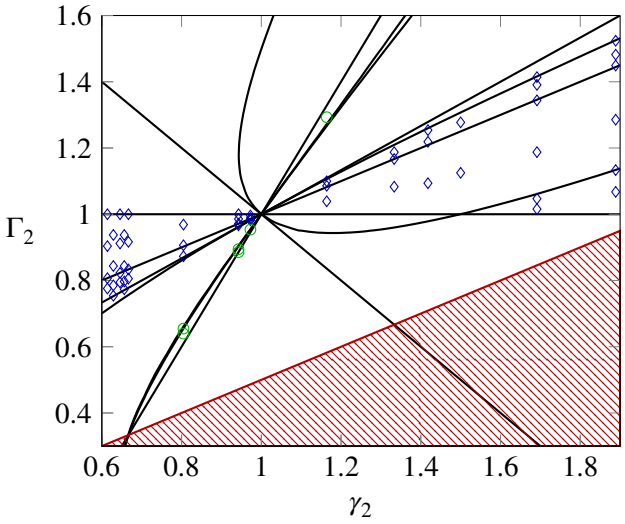

Figure 4.6: Decomposition of the parameter space near $(1, 1)$ (with the same conventions as in Figure 4.3).

4.4.2. The general case

The variety V and the semi-algebraic set $\mathcal{B}$ are defined as in the previous section. We normalize again by $\gamma_1 = 1$, we assume that $2\Gamma_1 \geq 1$, $2\Gamma_2 \geq \gamma_2 > 0$, $(\gamma_2, \Gamma_2) \neq (1, \Gamma_1)$, and that $\Gamma_1 \neq 1$, $\Gamma_2 \neq \gamma_2$ (case of water).

Theorem 4.8. *Splitting the subset of* $\mathbb{R}^3$ *defined by* $2\Gamma_2 > \gamma_2 > 0$ *and* $2\Gamma_1 > 1$ *into open subsets where the number of real singularities of* V *in the fibers is constant, can be done by cutting out* 12 *irreducible surfaces, consisting of* 5 *planes,* 3 *quadrics, two surfaces of degree 9 and one of degree 14.*

These polynomials were obtained by applying the algorithms from Section 4.3 to our system. The elimination steps were done using both Gröbner bases with FGb or with F5, and with triangular sets with RegularChains. Table 4.1 presents some timings for these methods (for computations done with interpolation, we give the results as $a \times b$ where a is the interpolation degree and b the time taken for each specialized computation).

Part III

Appendices

Appendix A

Source code

A.1. Magma source code for Matrix-F_5 (Chapters 2 and 3)

A.1.1. Algorithm Matrix-F_5

File `F5Mtx.m`

```
// load "F5_comptes_clean.m";

// *** Aux functions
// ****** General purpose

function Status (Calc)
    // String representing the current status of the calculation, in a
    // very short way. It is meant to be used in verbose output.
    res := Sprintf("[%o](n=%o)", Cputime(), Calc'n);
    return res;
end function;

function SortedMonomialsOfWeightedDegree(P,d)
    // Returns the sequence of monomials of W-deg d, sorted with
    // decreasing order
    M := MonomialsOfWeightedDegree(P,d);
    l := IndexedSetToSequence(M);
    Sort(~l); // Inc. order
    Reverse(~l); // Reversion for dec. order
    return l;
end function;

function DReg (F)
    // Computes the theoretical degree of regularity of F
    D := [Degree(f) : f in F];
    res := &+[d-1 : d in D] +1;
    return res;
end function;

function MinVariableDividing(mu)
```

```
    // Returns the smallest variable dividing monomial mu
    // /!\ Smallest variable = Highest index /!\
    P := Parent(mu);
    e := Exponents(mu);
    i := Max([i : i in [1..#e] | e[i] ne 0 ]);
    x := Parent(mu).i;
    return x;
end function;

function LeadingMonomialOrZero(f)
    // Same as LeadingMonomial, but returns 0 instead of an error
    // if f=0
    if f eq 0 then
        return 0;
    else
        return LeadingMonomial(f);
    end if;
end function;

function PolynomialToMatrixRow(f,Columns,N)
    P := Parent(f);
    R := BaseRing(P);
    Coefs,Mons := CoefficientsAndMonomials(f);
    M := ZeroMatrix(R,1,N);
    /* printf "Mons : %o\nColumns : %o\n\n", Mons, Columns; */
    for c in [1..#Coefs] do
        i := Columns[Mons[c]];
        M[1,i] := Coefs[c];
    end for;
    return M;
end function;

function MatrixRowToPolynomial(L,Monomials)
    P := Parent(Monomials[1]);
    res := P!0;
    for i in [1..NumberOfColumns(L)] do
        res +:= L[i]*Monomials[i];
    end for;
    return res;
end function;

function MtxInsertRow (M,L,l)
    M2 := VerticalJoin(M,L);
    return M2;
end function;

function MtxAddRow (M,c,i,j)
    // M[j] <- M[j] + c*M[i]
    M2 := AddRow(M,c,i,j);
```

```
    return M2;
end function;

function MtxMultiplyRow (M,c,i)
    // M[i] <- c*M[i]
    M2 := MultiplyRow(M,c,i);
    return M2;
end function;

procedure PrintDebug (s)
    if GetVerbose("Groebner") ge 3 then
        print s;
    end if;
end procedure;

function RescalePoly(f)
    if f eq 0 then
        return f;
    else
        c := LeadingCoefficient(f);
        f2 := (1/c)*f;
        return f2;
    end if;
end function;

// ****** Declarations and functions for signatures

// Record format for polynomials and their signatures
Sig := recformat
       <mu,   // The signature itself
        i : Integers(),    // The base polynomial
        row : Integers(), // The row in the matrix
        valbefore : RngMPolElt,
        valafter : RngMPolElt,
        d : Integers(),    // The degree of the associated GB poly
        log_redlist, // The list of the indices reducing this one
        LTbefore, // Leading term of the associated polynomial
        LTafter,
        LTfromGB
       >;

function SigCreate (i,f,row)
    P := Parent(f);
    s := rec <Sig |>;
    s`mu := P!1;
    s`i := i;
    s`row := row;
    s`valbefore := f;
    s`valafter := s`valbefore;
```

```
    s‘d := Degree(f);
    s‘log_redlist := [];
    s‘LTbefore := LeadingMonomialOrZero(f);
    s‘LTafter := s‘LTbefore;
    s‘LTfromGB := s‘LTbefore;
    return s;
end function;

function SigEqual (s1,s2)
    return (s1‘i eq s2‘i) and (s1‘mu eq s2‘mu);
end function;

function SigMultiplyByVar (s,v,row)
    // INPUT :
    // s : signature
    // v : variable
    // row : row
    // OUTPUT :
    // Signature of v*s, with row row
    PrintDebug(Sprintf("Multiplying by %o the signature %o\n",v,s));
    res := s;
    res‘mu := v*s‘mu;
    res‘valbefore := v*s‘valafter;
    res‘valafter := res‘valbefore;
    res‘log_redlist := [];
    res‘row := row;
    res‘LTbefore := v*s‘LTafter;
    res‘LTafter := res‘LTbefore;
    res‘d := 1 + s‘d;
    PrintDebug(Sprintf("Result : %o\n",res));
    return res;
end function;

function SigUpdate (s,Calc)
    // Update the value of s anywhere it appears in Calc
    mu := s‘mu;
    d := s‘d;
    i := s‘i;
    l := s‘row;
    Calc‘SigsFromMons[i][mu] := s;
    Calc‘SigsFromRows[d][l] := s;
    return Calc;
end function;

// ****** Specific aux functions

function F5_InsertRow(s,Calc)
    // Do what is necessary to insert the row defined by
    // signature s into the calculation Calc.
```

```
    // It means :
    // - inserting the row in the matrix
    // - inserting the signature in the list of indices
    // - updating the current and last row
    vprintf Groebner,3 :
    "Inserting the row defined by %o\ninto the calculation%o\n\n",
    s, Calc;
    f := s`valbefore;
    i := Calc`i_current;
    d := Calc`d_current;
    row := Calc`row_current;
    mu := s`mu;
    MatrixRow := PolynomialToMatrixRow(f,Calc`Columns[d],Calc`N[d]);
    Calc`M[d] := MtxInsertRow(Calc`M[d],MatrixRow,Calc`row_current);
    Calc`SigsFromMons[i][mu] := s;
    Calc`SigsFromRows[d][row] := s;
    Calc`lastrow[d][i] := Calc`lastrow[d][i] +1;
    Calc`row_current := Calc`row_current +1;
    col := Calc`Columns[d][s`LTbefore];
    if not IsDefined(Calc`Pivots[d],col) then
        Calc`Pivots[d][col] := row;
    end if;
    return Calc;
end function;

function F5_Criterion (s, Calc)
    // Returns true iff the signature s defines a polynomial
    // which should be added to the matrix.
    //
    // Criterion : true iff the monomial part of s is not a
    // leading term in the matrix at degree d-di and for the
    // i-1'th polynomial
    mu := s`mu;
    d_di := Degree(mu); // d_di = d_current - di
    i := Calc`i_current;
    try
        res := (mu notin Calc`Crit[d_di][i-1]);
    catch e
        res := true;
    end try;
    return res;
end function;

function F5_EchelonForm (Calc : Full := false)
    // Computes the row-echelon form as required by the F5 algorithm
    // (no row or column swapping)
    vprintf Groebner,3 : "Reduction to echelon form of matrix\n%o\n\n",
                         Calc`M[Calc`d_current];
```

```
    Comp := (assigned Calc‘OpCntRed);
    d := Calc‘d_current;
    i := Calc‘i_current;
    M := Calc‘M[d];
    N := Calc‘N[d];
    if i eq 1 then firstrow := 1;
    else firstrow := Calc‘lastrow[d][i-1] +1;
    end if;
    lastrow := Calc‘lastrow[d][i];
    Pivots := Calc‘Pivots[d];
    for row in [firstrow..lastrow] do
        s := Calc‘SigsFromRows[d][row];
        LTinit := s‘LTbefore;
        colinit := Calc‘Columns[d][LTinit];
        col := colinit;
        leadingcol := colinit;
        donetop := (row eq Pivots[colinit]);
        donefull := donetop;
        while (col le N) and (not donefull) do
            if (M[row][col] ne 0) then
                if (IsDefined(Pivots,col)) and Pivots[col] lt row then
                    Append(~s‘log_redlist,Pivots[col]);
                    M := MtxAddRow(M,-M[row][col],Pivots[col],row);
                    if Comp then
                        nop := NumberOfNonZeroEntries(M[Pivots[col]])
                               -1;
                        Calc‘OpCntRed +:= nop;
                    end if;
                elif not donetop then
                    leadingcol := col;
                    Pivots[leadingcol] := row;
                    donetop := true;
                    if not Full then
                        donefull := true;
                    end if;
                end if;
            end if;
            col +:= 1;
        end while;
        try
            M := MtxMultiplyRow(M,1/M[row][leadingcol],row);
        catch e;
        end try;
        s‘valafter := MatrixRowToPolynomial(M[row],Calc‘Monomials[d]);
        s‘LTafter := Calc‘Monomials[d][leadingcol];
        if s‘LTafter ne s‘LTbefore then
            s‘LTfromGB := s‘LTafter;
        end if;
        Calc := SigUpdate(s,Calc);
```

```
    end for;
    Calc'Pivots[d] := Pivots;
    Calc'M[d] := M;
    PrintDebug(Sprintf("Result :\n%o\n",M));
    return Calc;
end function;

function F5_GBCriterion (s,Calc)
    // INPUT :
    // - s : signature Calc : state of the computation
    // OUTPUT :
    // - res : boolean stating if we should add the
    // polynomial associated to that signature to the Groebner
    // basis
    res := (((assigned s'LTafter)
             and (s'LTafter ne s'LTbefore))
            or ((Calc'i_current eq 1)
                and (s'mu eq 1)));
    return res;
end function;

function F5_ReduceGBMatrix(Calc : Full := false)
    // Reduce the Groebner basis in the matrix
    Comp := (assigned Calc'OpCntRed);
    for d in [1..Calc'dreg] do
        M := Calc'M[d];
        firstrow := 1;
        lastrow := NumberOfRows(M);
        N := NumberOfColumns(M);
        Pivots := Calc'Pivots[d];

        for row in [firstrow..lastrow] do
            reduced_p := false;
            s := Calc'SigsFromRows[d][row];
            if s'LTafter ne s'LTbefore then // Element of the GB
                col := Calc'Columns[d][s'LTafter]+1;
                while (col le N) do
                    if (M[row][col] ne 0)
                       and IsDefined(Pivots, col)
                       and (Full or Pivots[col] gt row)
                       then
                        reduced_p := true;
                        M := MtxAddRow(M,-M[row][col],Pivots[col],row);
                        if Comp then
                            nop:=NumberOfNonZeroEntries(M[Pivots[col]])
                                   -1;
                            Calc'OpCntInterRed +:= nop;
                        end if;
                    end if;
```

```
                    col +:= 1;
                end while;
                // Update the value:
                s‘valafter := MatrixRowToPolynomial(M[row],
                                                    Calc‘Monomials[d]);
                Calc := SigUpdate(s,Calc);


                if reduced_p then
                    vprintf Groebner,2: "%o -> %o -> %o\n",
                                        s‘LTbefore, s‘LTafter,
                                        LeadingTerm(s‘valafter);
                end if;

            else // See if the original one was reduced
                sig := s‘LTafter div s‘LTfromGB;
                d0 := Degree(s‘LTfromGB);
                row0 := Calc‘Pivots[d0][Calc‘Columns[d0][s‘LTfromGB]];
                M[row] := PolynomialToMatrixRow(
                          MatrixRowToPolynomial(Calc‘M[d0][row0],
                                                Calc‘Monomials[d0])
                          *sig,
                          Calc‘Columns[d],N)[1];
            end if;
        end for;
        Calc‘M[d] := M;
    end for; // for d
    return Calc;
end function;

// *** Main function

// Record containing information about the F5 run. See the
// initialization step in function F5_Mtx for explanations of the
// fields.
Calculation := recformat<F,
                         m,
                         n,
                         dreg,
                         D,
                         quadratic,
                         LTGB,
                         NonCrit,
                         M,
                         Mtilde,
                         Monomials,
                         Columns,
                         Crit,
                         SigGB,
```

```
                        SigsFromMons,
                        SigsFromRows,
                        ReducedSigs,
                        Pivots,
                        N,
                        allMonomials,
                        lastrow,
                        i_current,
                        d_current,
                        row_current,
                        SigFromLT,
                        LTFromSig,
                        OpCntRed, OpCntInterRed, OpCntTot>;

function F5_Mtx (F : dreg := DReg(F),
                     bound := func<s | "">,
                     Full := false,
                     InterRed := false,
                     FullInterRed := Full and InterRed,
                     Comp := false)
    // bound : function taking the signature of a row and
    // printing various info about it
    //
    // Full : false -> no full reduction (only top)
    //        true  -> full reduction (for rows non-triv. top-reduced)
    //
    // InterRed : true iff we want to reduce the intermediate
    // Groebner bases
    //
    // FullInterRed : true iff we want to do this as a full
    // reduction
    //
    // Bases after completing step i

    // {{{ *** Initialization
    // {{{ ****** Preparation of input
    if exists(i){i : i in [1..#F] | not IsHomogeneous(F[i])} then
        error Sprintf("Polynomial #%o is not homogeneous", i);
    end if;

    P := Parent(F[1]);
    R := BaseRing(P);
    n := Rank(P);
    m := #F;
    D := [Degree(f) : f in F];
    // }}}
    // {{{ ****** Initialization of output
    G := AssociativeArray([0..m]); // Groebner bases
    G[0] := [];
```

```
    S := []; // Signatures
    // }}}
    // {{{ ****** Initialization of internal values

    Calc := rec<Calculation|>;
    // We will store the current state of the calculation there

    Calc`F := F;
    Calc`m := m;
    Calc`n := n;
    Calc`dreg := dreg;
    Calc`D := D;

    Calc`N := AssociativeArray([0..dreg]);
    // Calc`N[d] = number of monomials at degree d

    Calc`Monomials := AssociativeArray([0..dreg]);
    // Calc`Monomials[d][i] = i'th monomial of degree d

    Calc`Columns := AssociativeArray([0..dreg]);
    // Calc`Columns[d][m] = index of the column of monomial m at
    // degree d

    Calc`Pivots := AssociativeArray([0..dreg]);

    Calc`Crit := AssociativeArray([0..dreg]);
    // Calc`Crit[d] = sequence of monomials at degree d serving
    // in the F5 criterion

    Calc`lastrow := AssociativeArray([0..dreg]);
    // Calc`lastrow[d][i] = last row reached at the step i at degree d
    Calc`row_current := 1;

    Calc`M := AssociativeArray([0..dreg]);
    // Calc`M[d] = matrix built at degree d

    Calc`SigsFromMons := AssociativeArray([1..m]);
    Calc`SigsFromRows := AssociativeArray([0..dreg]);
    // Calc`SigsFromMons[i][mu] is the sig associated to the pair
    // i,mu
    //
    // Calc`SigsFromRows[d][l] is the sig associated to the row l
    // at degree d

    for d in [0..dreg] do
        Calc`Monomials[d] := SortedMonomialsOfWeightedDegree(P,d);
        Calc`N[d] := #(Calc`Monomials[d]);
        Calc`Pivots[d] := [];
        Calc`Columns[d] := AssociativeArray(Calc`Monomials[d]);
```

```
        for i := 1 to Calc`N[d] do
            Calc`Columns[d][Calc`Monomials[d][i]] := i;
        end for;

        Calc`Crit[d] := AssociativeArray([1..m]);

        Calc`lastrow[d] := [];

        Calc`M[d] := ZeroMatrix(R,0,Calc`N[d]);

        Calc`SigsFromRows[d] := AssociativeArray();
    end for;

    Calc`allMonomials := [];
    for d in [0..dreg] do
        Calc`allMonomials cat:= Calc`Monomials[d];
    end for;

    if Comp then
        Calc`OpCntRed := 0;
        Calc`OpCntInterRed := 0;
        Calc`OpCntTot := 0;
    end if;

    // }}}
    // }}}
    // {{{ *** Computation
    F := [RescalePoly(f) : f in F];

    for i in [1..m] do
        // {{{ ****** Preparation
        Calc`i_current := i;
        fi := F[i];
        di := D[i];
        Calc`SigsFromMons[i] := AssociativeArray(Calc`allMonomials);
        for d in [1..dreg] do
            if i eq 1 then
                Calc`lastrow[d][i] := 0;
                Calc`Crit[d][i] := [];
            else
                Calc`lastrow[d][i] := Calc`lastrow[d][i-1];
                Calc`Crit[d][i] := Calc`Crit[d][i-1];
            end if;
        end for;
        // }}}
        for d in [di..dreg] do
            vprintf User2 : "%o i=%o, d=%o\n", Status(Calc), i, d;
            // {{{ ****** Construction of the matrix
            Calc`d_current := d;
```

```
            Calc`row_current := Calc`lastrow[d][i]+1;
            if d eq di then
                // {{{ Case of a new polynomial in the matrix
                s := SigCreate(i,fi,Calc`row_current);
                Calc := F5_InsertRow(s, Calc);
                // }}}
            elif d gt di then
                // {{{ Case of a polynomial from a lower degree
                for k := Calc`N[d-di] to 1 by -1 do // Decreasing order
                    mu := Calc`Monomials[d-di][k];
                    PrintDebug(Sprintf("Considering signature %o",
                                       <i,mu>));
                    s := rec <Sig| i := i, mu := mu, d := d>;
                    if F5_Criterion(s, Calc) then
                        PrintDebug("F5 criterion : OK");
                        x := MinVariableDividing(mu);
                        mu2 := mu/x;
                        if IsDefined(Calc`SigsFromMons[i],mu2) then
                            // {{{ Construction and addition of
                            // the new poly
                            PrintDebug("Presence at "
                                       cat "previous degree : OK");
                            s2 := Calc`SigsFromMons[i][mu2];
                            s := SigMultiplyByVar(s2,x,
                                                  Calc`row_current);
                            Calc := F5_InsertRow(s, Calc);
                            // }}}
                        else
                            PrintDebug("Presence at "
                                       cat "previous degree : No");
                        end if;
                    else
                        PrintDebug("F5 criterion : No");
                    end if;
                end for;
                // }}}
            end if;
            // }}}
            // {{{ ****** Reduction of the matrix
            Calc := F5_EchelonForm(Calc : Full := Full);
            // }}}
        end for; // for d

        // {{{ 2.3. Update of the GB
        if InterRed then
            // "Reduce the Groebner basis", in matrix words
            Calc := F5_ReduceGBMatrix(Calc : Full := FullInterRed);

            // And recompute the Groebner basis
```

```
            jmin := 1;
            G[i] := [];
        else
            jmin := i;
            G[i] := G[i-1];
        end if;

        for j in [jmin..i] do
            for m in Keys(Calc`SigsFromMons[j]) do
                s := Calc`SigsFromMons[j][m];
                d := s`d;
                Append(~Calc`Crit[d][j],s`LTafter);
                if F5_GBCriterion(s,Calc) then
                    g := s`valafter;
                    Append(~G[i],g);
                end if;
                // }}}
            end for; // for m
        end for; // for j
    end for; // for i
    // }}}
    // {{{ *** Conclusion
    if Comp then
        Calc`OpCntTot := Calc`OpCntRed + Calc`OpCntInterRed;
    end if;
    return G, Calc;
    // }}}
end function;
// }}}


// Set of functions to help extracting information out of the Calc
// output of F5_Mtx

function FoldAlongGB (L)
    // Given a list of <s,list of whatever>, returns a array whose
    // entry m is <l,ls> where l is the cat of all lists associated to
    // a signature of leading term m and ls is the list of these
    // signatures.
    res := AssociativeArray();
    for k in Keys(L) do
        cpl := L[k];
        s := cpl[1];
        l := cpl[2];
        m := s`LTfromGB;
        if not IsDefined(res,m) then
            res[m] := <[],[]>;
        end if;
        rescpl := res[m];
```

```
        res[m] := <Append(rescpl[1],s),rescpl[2] cat cpl[2]>;
    end for;
    return res;
end function;

function ListReductionsDownwards (Calc)
   // Returns a list of <s,l> where s is a sig and l the list of
   // lines it reduced
    res := AssociativeArray();
    for i in [1..Calc`m] do
        SfM := Calc`SigsFromMons[i];
        SfL := Calc`SigsFromLines;
        for m in Keys(SfM) do
            s := SfM[m];
            L := s`log_redlist;
            for line in L do
                s2 := SfL[s`d][line];
                i2 := s2`i;
                mu2 := s2`mu;
                if not IsDefined(res,<i2,mu2>) then
                    res[<i2,mu2>] := <s2,[]>;
                end if;
                cpl := res[<i2,mu2>];
                l2 := Append(cpl[2],s);
                res[<i2,mu2>] := <cpl[1],l2>;
            end for;
        end for;
    end for;
    return res;
end function;
```

A.1.2. Weighted homogeneous systems

File `Whomo.m`

```
function TotalWeightedDegree (f)
    res := 0;
    for m in Monomials(f) do
        d := WeightedDegree(m);
        res := Max(res,d);
    end for;
    return res;
end function;


function Homogenize (F)
    if F eq [] then return [];
    else
        P := Parent(F[1]);
        n := Rank(P);
```

```
        ord := MonomialOrder(P);
        W := Grading(P);
        Wh := Append(W,1);
        if ord[1] eq "grevlexw" then
            nouvord := <"grevlexw",Wh>;
        else
            nouvord := ord;
        end if;
        R := CoefficientRing(P);
        Q := PolynomialRing(R,Wh,nouvord);
        Fh := [];
        for f in F do
            D := TotalWeightedDegree(f);
            fh := 0;
            M := Monomials(f);
            C := Coefficients(f);
            r := #M;
            for i in [1.. r] do
                e := Exponents(M[i]);
                d := &+[W[i] * e[i] : i in [1..n]];
                eh := Append(e,D - d);
                fh := fh + C[i] * Monomial(Q,eh);
            end for;
            Append(~Fh,fh);
        end for;
        return Fh,Q;
    end if;
end function;

function DeHomogenize (Fh,Ph,P)
    n := Rank(P);
    F := [];
    for f in Fh do
        C,M := CoefficientsAndMonomials(f);
        M2 := [];
        for i in [1..#M] do
            e := Exponents (M[i]);
            e2 := [];
            for k := 1 to n do // Remove the last variable
                e2[k] := e[k];
            end for;
            M2[i] := Monomial(P,e2);
        end for;
        f2 := Polynomial(C,M2);
        Append(~F,f2);
    end for;
    return F;
end function;
```

```
function WHomoToHomo (F)
    if F eq [] then return []; end if;
    P := Parent(F[1]);
    W := Grading(P);
    n := #W;
    R := CoefficientRing(P);
    Q := PolynomialRing(R,n,"grevlex");
    h := hom<P -> Q | [(Q.i)^W[i] : i in [1..n]]>;
    return h(F);
end function;

function FindWeights(F)
    // Find weights w1..wn such that all polynomials of F can be
    // written as polynomials in X1^w1..Xn^wn
    if F eq [] then error("Empty sequence"); end if;
    P := Parent(F[1]);
    n := Rank(P);
    W := [0 : i in [1..n]];
    for f in F do
        for m in Monomials(f) do
            e := Exponents(m);
            for i in [1..n] do
                W[i] := Gcd(W[i],e[i]);
            end for;
        end for;
    end for;
    return W;
end function;

function HomoToWHomo(F : W := [])
    if F eq [] then return []; end if;
    P := Parent(F[1]);
    n := Rank(P);
    if IsEmpty(W) then
        W := FindWeights(F);
    end if;
    assert #W eq n;
    Q := PolynomialRing(CoefficientRing(P),W);
    F2 := [];
    for i in [1..#F] do
        f := F[i];
        f2 := 0;
        C,M := CoefficientsAndMonomials(f);
        for j in [1..#C] do
            c := C[j];
            e := Exponents(M[j]);
            e2 := [e[k] div W[k] : k in [1..n]];
            m2 := Monomial(Q,e2);
            f2 +:= c*m2;
```

```
        end for;
        Append(~F2,f2);
    end for;
    return F2;
end function;
```

File `tests_GB.m`

```
load "F5Mtx.m";
load "Whomo.m";


SetVerbose("Groebner",1);
SetVerbose("User2",1);

W := [4,2,1,1];
P<X,Y,Z,T> := PolynomialRing(GF(65521),W);
F := [
    11541*X^2 + 53990*X*Y^2 + 37592*Y^4 + 27834*X*Y*Z^2 +
    30004*Y^3*Z^2 + 8314*X*Z^4 + 4825*Y^2*Z^4 + 28158*Y*Z^6 +
    19050*Z^8 + 26015*X*Y*Z*T + 557*Y^3*Z*T + 58180*X*Z^3*T +
    42158*Y^2*Z^3*T + 34612*Y*Z^5*T + 45718*Z^7*T + 34908*X*Y*T^2
    + 23734*Y^3*T^2 + 57434*X*Z^2*T^2 + 28743*Y^2*Z^2*T^2 +
    18580*Y*Z^4*T^2 + 10813*Z^6*T^2 + 23568*X*Z*T^3 +
    65399*Y^2*Z*T^3 + 3761*Y*Z^3*T^3 + 51403*Z^5*T^3 +
    20744*X*T^4 + 46116*Y^2*T^4 + 44410*Y*Z^2*T^4 + 37156*Z^4*T^4
    + 13625*Y*Z*T^5 + 24856*Z^3*T^5 + 39692*Y*T^6 + 34019*Z^2*T^6
    + 53446*Z*T^7 + 1929*T^8,

    33098*X^2 + 54375*X*Y^2 + 14826*Y^4 + 3377*X*Y*Z^2 +
    23825*Y^3*Z^2 + 7023*X*Z^4 + 52920*Y^2*Z^4 + 23589*Y*Z^6 +
    50207*Z^8 + 49154*X*Y*Z*T + 3423*Y^3*Z*T + 225*X*Z^3*T +
    43216*Y^2*Z^3*T + 56563*Y*Z^5*T + 8367*Z^7*T + 64946*X*Y*T^2
    + 60800*Y^3*T^2 + 65240*X*Z^2*T^2 + 34223*Y^2*Z^2*T^2 +
    39536*Y*Z^4*T^2 + 15290*Z^6*T^2 + 33901*X*Z*T^3 +
    21724*Y^2*Z*T^3 + 1521*Y*Z^3*T^3 + 32997*Z^5*T^3 +
    11568*X*T^4 + 35339*Y^2*T^4 + 39002*Y*Z^2*T^4 + 6645*Z^4*T^4
    + 38351*Y*Z*T^5 + 52984*Z^3*T^5 + 50226*Y*T^6 + 12349*Z^2*T^6
    + 50105*Z*T^7 + 8711*T^8,

    35779*X^2 + 7522*X*Y^2 + 49129*Y^4 + 6315*X*Y*Z^2 +
    33104*Y^3*Z^2 + 59779*X*Z^4 + 47533*Y^2*Z^4 + 14703*Y*Z^6 +
    51257*Z^8 + 11944*X*Y*Z*T + 27755*Y^3*Z*T + 55382*X*Z^3*T +
    1529*Y^2*Z^3*T + 24962*Y*Z^5*T + 6285*Z^7*T + 62689*X*Y*T^2 +
    61121*Y^3*T^2 + 2430*X*Z^2*T^2 + 52833*Y^2*Z^2*T^2 +
    48100*Y*Z^4*T^2 + 44590*Z^6*T^2 + 59954*X*Z*T^3 +
    43520*Y^2*Z*T^3 + 16159*Y*Z^3*T^3 + 112*Z^5*T^3 + 48989*X*T^4
    + 60810*Y^2*T^4 + 51752*Y*Z^2*T^4 + 54345*Z^4*T^4 +
    12978*Y*Z*T^5 + 14453*Z^3*T^5 + 20513*Y*T^6 + 46414*Z^2*T^6 +
    40677*Z*T^7 + 35597*T^8,
```

```
    59914*X^2 + 5160*X*Y^2 + 51256*Y^4 + 688*X*Y*Z^2 +
    37687*Y^3*Z^2 + 7159*X*Z^4 + 60215*Y^2*Z^4 + 39922*Y*Z^6 +
    11761*Z^8 + 22386*X*Y*Z*T + 45230*Y^3*Z*T + 28100*X*Z^3*T +
    27809*Y^2*Z^3*T + 54179*Y*Z^5*T + 18563*Z^7*T + 38052*X*Y*T^2
    + 64357*Y^3*T^2 + 26800*X*Z^2*T^2 + 37661*Y^2*Z^2*T^2 +
    57333*Y*Z^4*T^2 + 37124*Z^6*T^2 + 18538*X*Z*T^3 +
    29552*Y^2*Z*T^3 + 44263*Y*Z^3*T^3 + 17949*Z^5*T^3 +
    55615*X*T^4 + 32246*Y^2*T^4 + 61559*Y*Z^2*T^4 + 46600*Z^4*T^4
    + 59196*Y*Z*T^5 + 19035*Z^3*T^5 + 43459*Y*T^6 + 60433*Z^2*T^6
    + 10492*Z*T^7 + 57095*T^8
];

FH := WHomoToHomo(F);
//print FH;
assert IsHomogeneous(FH);

FF1 := HomoToWHomo(FH);
//print FF1;
assert [P!f : f in FF1] eq F;

FF2 := HomoToWHomo(FH : W := W);
//print FF2;
assert [P!f : f in FF2] eq F;

G1 := HomoToWHomo(GroebnerBasis(FH) : W := W);
G1 := [P!f : f in G1];
GG2 := F5_Mtx(F);
G2 := GG2[4];

SetVerbose("Groebner",0);
assert IsGroebner(G1);
assert IsGroebner(G2);
assert Ideal(G1) eq Ideal(F);
assert Ideal(G2) eq Ideal(F);
```

A.2. Maple source code for determinantal varieties (Chapter 4)

A.2.1. Algorithms

File `functions.mpl`

```
# read "functions.mpl";

with(LinearAlgebra):
with(FGb):
with(VectorCalculus):
with(combinat):

randomize():
```

```
randval := rand(100..10000):

Minors := proc(s,K)
description "Input: s integer, K matrix"
    "Output: the list of (s X s) minors of K";
local d1,d2,MIN,j,l,listofminors,subm:
    (d1,d2):=Dimension(K):
    if s > min(d1,d2) then
        printf("ERROR: INPUT s TOO LARGE\n");
        return;
    fi;

    MIN:=Matrix(binomial(d1,s),binomial(d2,s)):
    for j from 1 to binomial(d1,s) do
        for l from 1 to binomial(d2,s) do
            subm := SubMatrix(K,(choose([seq(i,i=1..d1)],s))[j],
                              (choose([seq(i,i=1..d2)],s))[l]);
            MIN[j,l]:=Determinant(subm):
        od:
    od:

    listofminors:=seq(seq(MIN[j,l],l=1..binomial(d2,s)),
                      j=1..binomial(d1,s)):
    return(listofminors):
end:

dimension := proc(sys, vars := indets(sys))
description "Find the dimension of the variety defined by the system "
    "by cutting it with random hyperplanes";
    local dim, hyp, randgen,hh, v, gg,sysCut,j, firstdone;

    dim := nops(vars):

    hyp := []:
    randgen := rand(-500..500):
    for j in seq(i,i=1..dim) do
        hh := randgen():
        for v in vars do
            hh := hh + randgen()*v:
        od;
        hyp := [op(hyp),hh]:
    od:

    sysCut := [op(sys),op(hyp)]:

    gg := [1]:
    firstdone := false;
    while gg = [1] and dim >= 0 do
        if firstdone then
```

```
            dim := dim -1:
            sysCut := sysCut[1..-2]:
        else
            firstdone := true:
        fi:
        printf("Trying dimension %d\n",dim):
        gg := fgb_gbasis(sysCut,65521,vars,[],{"index"=2000000}):
    od:
    return dim:
end:

sqfr := proc(f)
    description "Square-free reduction of f";
    return mul(ff,ff in map(x -> x[1],factors(f)[2])):
end:

GB_interpolate := proc(sys,charac,vars1,vars2,varinter,deg)
description "Use interpolation on varinter up to degree deg to compute "
    "a Groebner basis of sys for the order eliminating vars1";
local points, bases, v, i, gb, res, t0, safety, tgtdeg, m,
    bases_i, pol, frct;
    v := -1:
    points := []:
    bases := []:
    safety := 1:
    tgtdeg := 2*deg + 2 + 2*safety:
    for i from 1 to tgtdeg do
        printf("GB %a/%a: ", i,tgtdeg):
        t0 := time():
        while v = -1 or v in points do
            v := randval():
        od:
        gb := fgb_gbasis_elim(eval(sys,varinter=v),charac,vars1,vars2,
                              {"index"=20000000}):
        points := [op(points),v]:
        bases := [op(bases),gb]:
        printf("done (%a) [v=%a]\n", time()-t0,v):
    od:
    res := []:

    m := mul(varinter-v,v in points):

    for i from 1 to nops(bases[1]) do
        bases_i := map(x -> x[i]/lcoeff(x[i])
                       , bases):
        pol := CurveFitting[PolynomialInterpolation](points,
                                                     bases_i,
                                                     varinter):
        frct := ratrecon(pol,m,varinter,deg+safety,deg+safety):
```

```
        res := [op(res),numer(frct)]:
    od:
    return res:
end:

GB_find_deg := proc(sys,vars1,vars2,varinter,ntrials := 10)
description "Find the maximal degree in variable varinter of "
  "the polynomials in a Groebner basis of sys eliminating vars1";
local deg, gb, ss, i, v, t0;
    deg := -1:
    printf("Finding degree (%a tests): ", ntrials):
    i := 0:
    facts := []:
    while i < ntrials do
        ss := {}:
        for v in vars2 do
            if v <> varinter then
                ss := {op(ss), v=randval()}:
            fi:
        od:
        gb := fgb_gbasis_elim(eval(sys,ss),0,vars1,vars2):
        for j from 1 to nops(gb) do
            if i = 0 then
                facts := [op(facts),0]:
            fi:
            facts[j] := gcd(facts[j],gb[j]);
        od:
        deg := max(map(degree,gb,varinter),deg):
        printf("|"):
        i := i+1:
    od:
    printf("\n"):
    facts_prod := sqfr(mul(p,p in facts));
    return deg, facts_prod:
end:


Elimination_codim1 := proc(F,vars,params,{algo:=default})
description "Compute a polynomial whose zeroes cover the "
    "projection of V(F) on the parameter space"
    ""
    "Compute a system of generators for the elimination ideal "
    "of F obtained by eliminated the variables vars"
    "Uses the algorithm ‘algo‘ to compute the basis. Admissible "
    "values for ‘algo‘ are:"
    "- gb_direct : compute an elimination Groebner basis directly "
    "- gb_interp : compute an elimination Groebner basis using "
    "evaluation/interpolation"
    "Any other value defaults to gb_interp."
```

```
    ;
local var,deg_interp,G:
    if algo = gb_direct then
        ### Direct
        G := fgb_gbasis_elim(F,0,vars,params ,{"verb"=3}):
        return G[1];
    else
        ### Interpolation with first param
        var := params[1]:
        deg_interp, facts := GB_find_deg(F,vars,params,var):
        G := GB_interpolate(F,0,vars,params,var,deg_interp):
        return facts*G[1];
    fi:
end:

DiscriminatingPolynomial := proc(M,r0,H,vars,params,
                                 eqs := [], neqs := [], {algo:=default})
description "Compute a polynomial P as in section 3."
    ""
    "The argument ‘eqs‘ is a set of extra equations and inequations "
    "restricting the solutions: the solutions returned are "
    "projections in the parameter space of points at which all "
    "polynomials of eqs vanish."
    ""
    "The polynomials in ‘neqs‘ are saturated in the computations."
    ""
    "See Elimination for the parameter ‘algo‘":
local res,res1,res2,res3,t0,t1,t2,t3:
    t0 := time():
    res1 := RankExactly(M,r0,vars,params,eqs,neqs,
                        ':-algo'=algo):
    t1 := time():
    res2 := DeterminantCritVals(M,r0,vars,params,eqs,neqs,true,
                                ':-algo'=algo):
    t2 := time():
    res3 := DeterminantBoundary(M,r0,H,vars,params,eqs,neqs,true,
                                ':-algo'=algo):
    t3 := time():
    printf("RankExactly\t%as\n",t1-t0):
    printf("DeterminantCritVals\t%as\n",t2-t1):
    printf("DeterminantBoundary\t%as\n",t3-t2):
    return res1*res2*res3, res1, res2, res3;
end:


IncidenceVariety := proc(M,r)
description "Compute a system of generators for the incidence variety "
    "of rank r of M"
    ;
```

```
local MatrixU, MatrixY,k,i,j,Prod1,Sys1,Prod2,Sys2,Sys:
    k := RowDimension(M):
    MatrixU := Matrix(k-r,k):
    for i from 1 to k-r do
        for j from 1 to k do
            MatrixU[i,j] := randval():
        od:
    od:

    print(MatrixU);

    MatrixY := Matrix(k,k-r):
    for i from 1 to k do
        for j from 1 to k-r do
            MatrixY[i,j] := Y[i,j]:
        od:
    od:
    Prod1 := M . MatrixY:
    Sys1 := [seq(seq(Prod1[i][j],j=1..k-r),i=1..k)]:
    Prod2 := MatrixU . MatrixY - Matrix(k-r,k-r,shape=identity):
    Sys2 := [seq(seq(Prod2[i][j],j=1..k-r),i=1..k-r)]:
    Sys := [op(Sys1),op(Sys2)]:
    return Sys, indets(MatrixY):
end:

RankExactly := proc(M,r0,vars,params,eqs := [], neqs := [],
                    {algo := default})
description "Algorithm RankExactly (section 3.4)"
    ""
    "See DiscriminatingPolynomial for a description of parameters "
    "‘eqs‘ and ‘neqs‘."
    "See Elimination for a description of the parameter ‘algo‘";
local res,k,n,FVr,dim,JVr,FV,F0,varsY,Mins,i,Sysi,G, codim:
    k := RowDimension(M):
    n := nops(vars):
    res := 1:

    FVr := [Minors(r0+1,M)]:
    codim := (k-r0)^2:
    JVr := Jacobian(FVr,[op(vars)]):
    FV := [op(FVr), Minors(codim,JVr),op(eqs),u*mul(f, f in neqs)-1]:
    F0, varsY := IncidenceVariety(M,r0):

    Mins := [Minors(r0,M)]:
    for i from 1 to nops(Mins) do
        Sysi := [op(FV),op(F0),op(Mins[1..i-1]),uu*Mins[i]-1]:
        g := Elimination_codim1(Sysi,[uu,u,op(varsY),op(vars)],params,
                                ':-algo'=algo):
        if g <> FAIL then
```

```
            res := res*g:
        fi:
    od:
    printf("RankExactly: %a\n", factor(res));
    return res:
end:

DeterminantCritVals := proc(M,r0,vars,params,eqs := [], neqs := [],
                            skipRankExactly := false, {algo:=default})
    description "Algorithm DeterminantCritVals (section 3.5)"
    ""
    "See DiscriminatingPolynomial for a description of parameters "
    "‘eqs‘ and ‘neqs‘."
    "See Elimination for a description of the parameter ‘algo‘":
local res,F0,varsY,N,J,F1,G:
    if not skipRankExactly then
        res := RankExactly(M,r0,vars,params):
    else
        res := 1:
    fi:
    F0, varsY := IncidenceVariety(M,r0-1):
    N := nops(F0):
    J := Jacobian(F0,[op(varsY),op(vars)]):
    F1 := [op(F0),Minors(N,J),op(eqs),u*mul(f, f in neqs)-1]:
    g := Elimination_codim1(F1,[u,op(varsY),op(vars)],params,
                            ':-algo'=algo):
    res := res*g:
    printf("CritVals: %a\n", factor(res));
    return res:
end:

DeterminantBoundary := proc(M,r0,H,vars,params,
                            eqs := [], neqs := [],
                            skipRankExactly := false,
                            {algo := default})
    description "Algorithm DeterminantBoundary (section 3.6)"
    ""
    "See DiscriminatingPolynomial for a description of parameters "
    "‘eqs‘ and ‘neqs‘."
    "See Elimination for a description of the parameter ‘algo‘";
local res, F0, varsY, h, F1, G:
    if not skipRankExactly then
        res := RankExactly(M,r0,vars,params):
    else
        res := 1:
    fi:
    F0, varsY := IncidenceVariety(M,r0-1):
    for h in H do
        F1 := [op(F0),h,op(eqs),u*mul(f, f in neqs)-1]:
```

```
        g := Elimination_codim1(F1,[u,op(varsY),op(vars)],params,
                                ':-algo'=algo):
        res := res * g:
    od:
    printf("Boundary: %a\n", factor(res));
    return res:
end:
```

A.2.2. Case of water

Definitions

File `water_def.mpl`

```
# read "water_def.mpl":

d1:=g1-G1:d2:=g2-G2:

MatrixD:=Matrix([[-G1*y1, -z1-1,-G1+d1*z1,2*d1*y1],
                 [-g1*z1,y1,d1*y1,(G1-d1)-2*d1*z1],
                 [-G2*y2, -z2-1,-G2+d2*z2,2*d2*y2],
                 [-g2*z2,y2,d2*y2,(G2-d2)-2*d2*z2]]):

g1 := 1: G1 := 1:

vars := [y1,y2,z1,z2]:
params := [g2,G2]:

H := [1-y1^2-(z1+1)^2, 1-y2^2-(z2+1)^2]:

k := 4: n := 4: t := 2:

DetD:=Determinant(MatrixD);
```

Computations using the functions

File `water_functions.mpl`

```
# read "water_functions.mpl":

# Function definitions

read "functions.mpl";


# Definitions for water

read "water_def.mpl";


# Computations
```

```
t0 := time():
res_general,res1,res2,res3 :=
    DiscriminatingPolynomial(MatrixD,3,H,vars,params,[],[],
                             algo = gb_direct):
t1 := time():


# Results

# RankExactly
ff1 := map(x -> x[1],factors(res1)[2]):
nops(ff1);
# 6
map(degree,ff1);
# [1,1,1,2,1,1]

# Critical values
ff2 := map(x -> x[1],factors(res2)[2]):
nops(ff2);
# 5
map(degree,ff2);
# [1, 1, 1, 1, 1]

# Boundary
ff3 := map(x -> x[1],factors(res3)[2]):
nops(ff3);
# 7
map(degree,ff3);
# [1, 1, 1, 1, 2, 2, 3]

allfacts := [op({op(ff1),op(ff2),op(ff3)})]:
nops(allfacts);
# 13
map(degree,allfacts);
# [1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 3]

filename := "water_results.mpl";
fid := fopen(filename,WRITE):
fprintf(fid,"## RankExactly:\n##---------\n"):
fprintf(fid,"res1 := [\n"):
for i from 1 to nops(ff1)-1 do
    f := ff1[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff1[-1]):
fprintf(fid,"## CritVals:\n##---------\n"):
fprintf(fid,"res2 := [\n"):
for i from 1 to nops(ff2)-1 do
```

```
    f := ff2[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff2[-1]):
fprintf(fid,"## Boundary\n##---------\n"):
fprintf(fid,"res3 := [\n"):
for i from 1 to nops(ff3)-1 do
    f := ff3[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff3[-1]):
fprintf(fid,"## All (without duplicates)\n##---------\n"):
fprintf(fid,"allfacts := [\n"):
for i from 1 to nops(allfacts)-1 do
    f := allfacts[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",allfacts[-1]):
fclose(fid):
```

Step-by-step computations

File `water_computations.mpl`

```
# read "water_computations.mpl";

read "functions.mpl";


# System

read "water_def.mpl":
DetD:=Determinant(MatrixD):
gradD_yz := [diff(DetD, y1), diff(DetD, z1),
             diff(DetD, y2), diff(DetD, z2)]:

r0 := 3:

Sys := [DetD,op(gradD_yz)]:


# Incidence varieties

F2,varsY2 := IncidenceVariety(MatrixD,2):
F3,varsY3 := IncidenceVariety(MatrixD,3):


# Singular points and critical values
# Rank(M) <= 2
```

```
Jac := Jacobian(F2,[z1,z2,y1,y2,op(varsY2)]):
Sys_rk2 := [op(F2),Minors(2*k+(k-r0+1)^2,Jac),u*g2-1]:

GB_rk2 := fgb_gbasis_elim(Sys_rk2,0,
                          [u,op(varsY2),z1,z2,y1,y2],[G2,g2]):
lprint(factor(GB_rk2));

(* Results:
[(G2-1)^2*(-2*g2+1+G2)^2*(-g2-1+2*G2)^2*(2*G2-g2)^2]
*)


# Intersection with the boundary

# Side 1
Sys_bnd1 := [op(F2),H[1]]:
GB_bnd1 := fgb_gbasis_elim(Sys_bnd1,0,
                           [op(varsY2),z1,z2,y1,y2],[G2,g2]):
lprint(factor(GB_bnd1));

(* Results:
[g2*(G2-1)^2*(-2*g2-1+3*G2)*(3*G2^2-5*G2*g2+g2^2+2*G2-2*g2+1)]
*)

# Side 2

Sys_bnd2 := [op(F2),H[2]]:
GB_bnd2 := fgb_gbasis_elim(Sys_bnd2,0,
                           [op(varsY2),z1,z2,y1,y2],[G2,g2]):
lprint(factor(GB_bnd2));
(* Results:
[(G2-1)^2*(2*G2^2-5*G2*g2+2*g2^2-2*G2+3*g2)\
 *(2*G2^3-3*G2^2*g2-3*G2*g2^2+2*g2^3+2*G2^2+9*G2*g2-11*g2^2-4*G2+6*g2)\
 *(-g2+2*G2)^2]
*)



# Rank(M) = 3

allGB_rk3 := []:
MM := [Minors(k-1,MatrixD)]:

Sys_rk3 := [DetD,op(gradD_yz),op(F3),u1*g2-1]:
for i from 1 to k^2 do
    Sys_rk3_i := [op(Sys_rk3),op(MM[1..i-1]),u2*MM[i]-1]:
    GB := fgb_gbasis_elim(Sys_rk3_i,0,
                          [u1,u2,op(varsY3),z1,z2,y1,y2],
                          [G2,g2],
```

```
                          {"verb"=3}):
    allGB_rk3 := [op(allGB_rk3),[i,GB]]:
od:

for g in allGB_rk3 do
    if g[2] <> [1] then
        printf("i%a Basis:%a\n",g[1],factor(g[2])):
    fi:
od:
(* Results:
i=1 Basis:[(-g2+2*G2)*(g2-2+G2)*(2*G2^2-5*G2*g2+2*g2^2+1)]
*)
```

Cylindrical algebraic decomposition

File `water_cad.mpl`

```
# restart; read "water_cad.mpl";

read "water_def.mpl";

SDetD := [DetD,diff(DetD,y1),diff(DetD,y2),
          diff(DetD,z1),diff(DetD,z2)]:

Pols := [g2,
         G2,
         g2-2*G2,
         G2-1,
         3*G2-1-2*g2,
         3*G2^2-5*G2*g2+g2^2+2*G2-2*g2+1,
         2*G2^2-5*G2*g2+2*g2^2-2*G2+3*g2,
         (2*G2^3-3*G2^2*g2-3*G2*g2^2+2*g2^3
          +2*G2^2+9*G2*g2-11*g2^2-4*G2+6*g2),
         -2*g2+G2+1,
         2*G2-1-g2,
         G2-2+g2,
         2*G2^2-5*G2*g2+2*g2^2+1];

xi := product(Pols[i], i = 1 .. nops(Pols));

with(RegularChains); with(ChainTools); with(SemiAlgebraicSetTools);
R := PolynomialRing([G2,g2]);
cadfull := CylindricalAlgebraicDecompose(Pols,R,output=cadcell):
sols := [];
for j from 1 to nops(cadfull) do
    sols := [op(sols),
             subs(op(2,op(1,subs(op(2,op(1,cadfull[j])),SamplePoint))),
                  box_bwe)[1]
            ]:
od:
```

```
print(nops(sols));
print(sols[1]);

avg2 := proc(l):
    return (l[1]+l[2])/2;
end:



insideBB := x -> (evala(eval(y1^2+(z1+1)^2-1, x)) <= 0 and
                  evala(eval(y2^2+(z2+1)^2-1, x)) <= 0):

PolsRed := remove(x -> x = g2 or x = G2 or x = 2*G2-g2, Pols):

sAvg := map(x -> [g2 = avg2(eval(g2,x)),
                        G2 = avg2(eval(G2,x))],
                  sols):

sAvgValid := select(x ->(eval(xi,x) <> 0
                                    and 0 < eval(g2,x)
                                    and 0 < eval(G2,x)
                                    and eval(g2,x) < 2*eval(G2,x)),
                               sAvg):
print(nops(%)):
# 570



Sols := [seq(select(x -> x <> {y1 = 0, y2 = 0,
                                z1 = -1, z2 = -1},
                    [solve(Groebner[Basis](eval(SDetD,
                                                sAvgValid[i]),
                                           plex(y1, y2, z1, z2)))]),
             i = 1..nops(sAvgValid))]:
print(map(nops, Sols));
# [2..2]

SolsInBB := [seq(select(x -> (evalb(0 <= evala(eval(y1^2, x)))
                              and evalb(0 <= evala(eval(y2^2, x)))
                              and insideBB(x)), Sols[i]),
                 i = 1 .. nops(Sols))]:

print(nops(select(x -> x <> [], SolsInBB)));
# 187

SolsBBSym := [seq(select(x -> (evalb(evala(eval(y1, x)) = 0)
                               and evalb(evala(eval(y2, x)) = 0)
```

```
                                and insideBB(x)), Sols[i]),
                  i = 1 .. nops(Sols))]:

print(nops(select(x -> ( x <> []), SolsBBSym)));
# 156

IndBBSym := select(x -> ( SolsBBSym[x] <> []),
                   [seq(i, i = 1 .. nops(SolsBBSym))]):

SolsBBNoSym := [seq(select(x -> (evalb(0 < evala(eval(y1^2, x)))
                                 and evalb(0 < evala(eval(y2^2, x)))
                                 and insideBB(x)), Sols[i]),
                    i = 1 .. nops(Sols))]:

print(nops(select(x -> ( x <> []), SolsBBNoSym)));
# 31

IndBBNoSym := select(x -> ( SolsBBNoSym[x] <> []),
                     [seq(i, i = 1 .. nops(SolsBBNoSym))]):

g2G2_1 := sAvgValid[IndBBSym]:
g2G2_2 := sAvgValid[IndBBNoSym]:

filename1 := "pts_cad_1.txt";
fid1 := fopen(filename1,WRITE);
for s in g2G2_1 do
    fprintf(fid1,"%a %a\n",evalf(eval(g2,s)), evalf(eval(G2,s)));
od:
fclose(fid1);

filename2 := "pts_cad_2.txt";
fid2 := fopen(filename2,WRITE);
for s in g2G2_2 do
    fprintf(fid2,"%a %a\n",evalf(eval(g2,s)), evalf(eval(G2,s)));
od:
fclose(fid2);

f0,f1,f2,f3,f4,f5,f6,f7,f8,f9 := op(Pols[3..-1]):

test_all := x -> map(f -> evalb(eval(f,x) > 0),
                     [f1,f2,f3,f4,f5,f6,f7,f8,f9]);




crit_1 := x -> ((eval(f1,x) > 0
                 and eval(f4,x) > 0
                 and eval(f2,x) < 0)
                or (eval(f1,x) < 0
```

```
                    and eval(f2,x) > 0));
print({op(map(crit_1,g2G2_1))});

print(nops(g2G2_1));
# 156
print(nops(select(crit_1,sAvgValid)));
# 156



crit_2 := x -> ((eval(f1,x) < 0
                 and eval(f6,x) > 0
                 and eval(f3,x) < 0)
                or (eval(f1,x) > 0
                    and eval(f6,x) < 0
                    and eval(f5,x) > 0)):
print({op(map(crit_2,g2G2_2))});

print(nops(g2G2_2));
# 31
print(nops(select(crit_2,sAvgValid)));
# 31
```

A.2.3. General case

Definitions

File `general_def.mpl`

```
# read "general_def.mpl";

d1:=g1-G1:d2:=g2-G2:

MatrixD:=Matrix([[-G1*y1, -z1-1,-G1+d1*z1,2*d1*y1],
                 [-g1*z1,y1,d1*y1,(G1-d1)-2*d1*z1],
                 [-G2*y2, -z2-1,-G2+d2*z2,2*d2*y2],
                 [-g2*z2,y2,d2*y2,(G2-d2)-2*d2*z2]]):

g1 := 1:
vars := [y1,y2,z1,z2]:
params := [G1,g2,G2]:

H := [1-y1^2-(z1+1)^2, 1-y2^2-(z2+1)^2]:

k := 4: n := 4: t := 3:

DetD:=Determinant(MatrixD);
```

Using the functions

File general_functions.mpl

```
# read "general_functions.mpl":

# Function definitions

read "functions.mpl";


# Definitions

read "general_def.mpl";


# Computations

t0 := time():
res_general,res1,res2,res3 :=
    DiscriminatingPolynomial(MatrixD,3,H,vars,params,[],[]):
t1 := time():


# Results

# RankExactly
ff1 := map(x -> x[1],factors(res1)[2]):
nops(ff1);
# 6
map(degree,ff1);
# [1,1,1,1,2,1]

# Critical values
ff2 := map(x -> x[1],factors(res2)[2]):
nops(ff2);
# 7
map(degree,ff2);
# [1, 1, 1, 14, 1, 1, 1]

# Boundary
ff3 := map(x -> x[1],factors(res3)[2]):
nops(ff3);
# 7
map(degree,ff3);
# [2, 9, 9, 2, 1, 1, 1]

allfacts := [op({op(ff1),op(ff2),op(ff3)})]:
nops(allfacts);
# 15
map(degree,allfacts);
# [1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 9, 9, 14]
```

```
filename := "general_results.mpl";
fid := fopen(filename,WRITE):
fprintf(fid,"## RankExactly:\n##---------\n"):
fprintf(fid,"res1 := [\n"):
for i from 1 to nops(ff1)-1 do
    f := ff1[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff1[-1]):
fprintf(fid,"## CritVals:\n##---------\n"):
fprintf(fid,"res2 := [\n"):
for i from 1 to nops(ff2)-1 do
    f := ff2[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff2[-1]):
fprintf(fid,"## Boundary\n##---------\n"):
fprintf(fid,"res3 := [\n"):
for i from 1 to nops(ff3)-1 do
    f := ff3[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",ff3[-1]):
fprintf(fid,"## All (without duplicates)\n##---------\n"):
fprintf(fid,"allfacts := [\n"):
for i from 1 to nops(allfacts)-1 do
    f := allfacts[i]:
    fprintf(fid,"    %a,\n", f):
od:
fprintf(fid,"    %a\n]:\n\n",allfacts[-1]):
fclose(fid):
```

Appendix B

Bibliography

[Agn02] G. Agnarsson. "On the Sylvester denumerants for general restricted partitions". In: *Proceedings of the Thirty-third Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 2002)*. Vol. 154. 2002, pp. 49–60 (cit. on p. 29).

[Alf05] J. L. R. Alfonsín. *The Diophantine Frobenius Problem*. Oxford Lecture Series in Mathematics and Its Applications. Oxford: Oxford University Press, 2005 (cit. on pp. 29, 30, 91).

[Alf98] J. L. R. Alfonsín. "On variations of the subset sum problem". In: *Discrete Applied Mathematics* 81.1–3 (1998), pp. 1–7. ISSN: 0166-218X. DOI: *10.1016/S0166-218X(96)00105-9* (cit. on p. 91).

[ALM99] P. Aubry, D. Lazard, and M. Moreno Maza. "On the theories of triangular sets". In: *Journal of Symbolic Computation, Special Issue on Polynomial Elimination* 28 (1999), pp. 105–124. DOI: *10.1006/jsco.1999.0269* (cit. on p. 8).

[Bar+05] M. Bardet, J.-C. Faugère, B. Salvy, and B.-Y. Yang. "Asymptotic Behaviour of the Degree of Regularity of Semi-Regular Polynomial Systems". In: *MEGA'05, 2005. Eighth International Symposium on Effective Methods in Algebraic Geometry*. 2005 (cit. on pp. 84, 87, 110).

[Bar04] M. Bardet. "Étude des systèmes algébriques surdéterminés. Applications aux codes correcteurs et à la cryptographie". Français. PhD thesis. Université Pierre et Marie Curie - Paris VI, Dec. 2004. URL: *http://tel.archives-ouvertes.fr/tel-00449609* (cit. on pp. 32, 33, 41, 47, 72, 87, 110–112).

[BCR98] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*. Vol. 36. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Translated from the 1987 French original, Revised by the authors. Springer-Verlag, Berlin, 1998, pp. x+430. ISBN: 3-540-64663-9. DOI: *10.1007/978-3-662-03718-8* (cit. on p. 135).

[BD07] C. W. Brown and J. H. Davenport. "The complexity of quantifier elimination and cylindrical algebraic decomposition". In: *ISSAC 2007*. ACM, New York, 2007, pp. 54–60. DOI: *10.1145/1277548.1277557* (cit. on p. 131).

[BFS04] M. Bardet, J.-C. Faugère, and B. Salvy. "On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations". In: *International Conference on Polynomial System Solving - ICPSS*. Paris, France, Nov. 2004, pp. 71–75. URL: *http://www-salsa.lip6.fr/~jcf/Papers/43BF.pdf* (cit. on p. 15).

[BFS14] M. Bardet, J.-C. Faugère, and B. Salvy. "On the Complexity of the F5 Gröbner basis Algorithm". In: *Journal of Symbolic Computation* (Sept. 2014), pp. 1–24. DOI: *10.1016/j.jsc.2014.09.025* (cit. on pp. 13, 15, 35, 60, 67–69, 76–79).

[Bon+12] B. Bonnard et al. "Geometric optimal control of the contrast imaging problem in nuclear magnetic resonance". In: *IEEE Trans. Automat. Control* 57.8 (2012), pp. 1957–1969. ISSN: 0018-9286. DOI: *10.1109/TAC.2012.2195859* (cit. on pp. 130, 133).

[Bon+13] B. Bonnard, M. Chyba, A. Jacquemard, and J. Marriott. "Algebraic geometric classification of the singular flow in the contrast imaging problem in nuclear magnetic resonance". In: *Math. Control Relat. Fields* 3.4 (2013), pp. 397–432. ISSN: 2156-8472. DOI: *10.3934/mcrf.2013.3.397* (cit. on pp. 19, 130, 131, 133, 142).

[Bon+16] B. Bonnard et al. "Determinantal sets, singularities and application to optimal control in medical imagery". In: *Proceedings of the 2016 International Symposium on Symbolic and Algebraic Computation*. ISSAC '16. Waterloo, Canada, 2016 (cit. on p. 129).

[BP99] M. de Boer and R. Pellikaan. "Gröbner bases for codes". In: *Some tapas in computer algebra*. Ed. by A. M. Cohen, H. Cuypers, and H. Sterk. Algorithms and Computation in Mathematics 4. Springer, 1999, pp. 237–259. URL: *http://www.win.tue.nl/%5C~%7B%7Druudp/paper/34.pdf* (cit. on p. 87).

[BS88] D. Bayer and M. Stillman. "On the complexity of computing syzygies". In: *Journal of Symbolic Computation* 6.2-3 (Oct. 1988), pp. 135–147. ISSN: 07477171. DOI: *10.1016/S0747-7171(88)80039-7* (cit. on p. 76).

[Buc76] B. Buchberger. "A theoretical basis for the reduction of polynomials to canonical forms". In: *ACM SIGSAM Bulletin* 10.3 (1976), pp. 19–29 (cit. on pp. 8, 65, 66).

[Buc79] B. Buchberger. "A Criterion for Detecting Unnecessary Reductions in the Construction of Gröbner Bases". In: 72 (Jan. 1979), pp. 3–21. DOI: *10.1007/3-540-09519-5* (cit. on p. 67).

[BV88] W. Bruns and U. Vetter. *Determinantal rings*. Vol. 1327. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1988, pp. viii+236. ISBN: 3-540-19468-1 (cit. on pp. 10, 14, 51).

[BW93] T. Becker and V. Weispfenning. *Gröbner bases*. Vol. 141. Graduate Texts in Mathematics. A computational approach to commutative algebra, In cooperation with Heinz Kredel. New York: Springer-Verlag, 1993, pp. xxii+574. ISBN: 0-387-97971-9. DOI: *10.1007/978-1-4612-0913-3* (cit. on p. 61).

[CD05] E. Cattani and A. Dickenstein. "Introduction to residues and resultants". In: *Solving polynomial equations*. Springer Berlin Heidelberg, 2005, pp. 1–61 (cit. on p. 8).

[CDR96] M. Caboara, G. de Dominicis, and L. Robbiano. "Multigraded Hilbert Functions and Buchberger Algorithm". In: *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation, ISSAC '96, Zurich, Switzerland, July 24-26, 1996*. Ed. by E. Engeler, B. F. Caviness, and Y. N. Lakshman. ACM, 1996, pp. 72–78. ISBN: 0-89791-796-0. DOI: *10.1145/236869.236901* (cit. on p. 84).

[Chy+15] M. Chyba et al. "Optimal Geometric Control Applied to the Protein Misfolding Cyclic Amplification Process". English. In: *Acta Applicandae Mathematicae* 135.1 (2015), pp. 145–173. ISSN: 0167-8019. DOI: *10.1007/s10440-014-9950-8* (cit. on p. 131).

[CKM97] S. Collart, M. Kalkbrenner, and D. Mall. "Converting bases with the Gröbner walk". English. In: *Journal of Symbolic Computation*. Special issue on computational algebra and number theory: proceedings of the first MAGMA conference 24.3-4 (1997), pp. 465–469. DOI: *10.1006/jsco.1996.0145* (cit. on p. 12).

[CLO07] D. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms*. Third. Undergraduate Texts in Mathematics. An introduction to computational algebraic geometry and commutative algebra. New York: Springer, 2007, pp. xvi+551. ISBN: 978-0-387-35650-1; 0-387-35650-9. DOI: *10.1007/978-0-387-35651-8* (cit. on pp. 27, 38, 63–66).

[Col75] G. E. Collins. "Quantifier elimination for real closed fields by cylindrical algebraic decomposition". In: *Automata theory and formal languages (Second GI Conf., Kaiserslautern, 1975)*. Vol. 33. Lecture Notes in Computer Science. Springer, Berlin, 1975, pp. 134–183 (cit. on pp. 8, 131).

[Com74] L. Comtet. *Advanced combinatorics*. enlarged. The art of finite and infinite expansions. D. Reidel Publishing Co., Dordrecht, 1974, pp. xi+343. ISBN: 90-277-0441-4 (cit. on pp. 29, 30).

[CS95] M. Coste and M. Shiota. "Thom's first isotopy lemma: a semialgebraic version, with uniform bound". In: *Real analytic and algebraic geometry (Trento, 1992)*. de Gruyter, Berlin, 1995, pp. 83–101 (cit. on p. 45).

[Demoo] M. Demazure. *Bifurcations and catastrophes*. Universitext. Geometry of solutions to nonlinear problems, Translated from the 1989 French original by D. Chillingworth. Springer-Verlag, Berlin, 2000, pp. viii+303. ISBN: 3-540-52118-6. DOI: *10.1007/978-3-642-57134-3* (cit. on p. 44).

[DH88] J. H. Davenport and J. Heintz. "Real quantifier elimination is doubly exponential". In: *Journal of Symbolic Computation* 5.1-2 (1988), pp. 29–35. ISSN: 0747-7171. DOI: *10.1016/S0747-7171(88)80004-X* (cit. on p. 131).

[DS06] G. Dalzotto and E. Sbarra. "Computations in weighted polynomial rings". In: *Analele Stiintifice ale Universitatii Ovidius Constanta* 14(2) (2006), pp. 31–44. URL: *http://www.anstuocmath.ro/mathematics/pdf12/31_44_GDalzotto_ESbarra.pdf* (cit. on p. 85).

[Eis95] D. Eisenbud. *Commutative algebra*. Vol. 150. Graduate Texts in Mathematics. With a view toward algebraic geometry. New York: Springer-Verlag, 1995, pp. xvi+785. ISBN: 0-387-94268-8; 0-387-94269-6. DOI: *10.1007/978-1-4612-5350-1* (cit. on pp. 25, 27, 30, 31, 43, 48, 49, 51, 85, 100).

[Fau+13] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. "Using Symmetries in the Index Calculus for Elliptic Curves Discrete Logarithm". English. In: *Journal of Cryptology* (2013), pp. 1–41. ISSN: 0933-2790. DOI: *10.1007/s00145-013-9158-5* (cit. on pp. 9, 18, 87, 120).

[Fau+14] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. "Sub-cubic change of ordering for Gröbner basis". In: *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation - ISSAC '14*. Kobe, Japan: ACM Press, July 2014, pp. 170–177. ISBN: 9781450325011. DOI: *10.1145/2608628.2608669* (cit. on pp. 80, 118).

[Fau+93] J.-C. Faugère, P. Gianni, D. Lazard, and T. Mora. "Efficient computation of zero-dimensional Gröbner bases by change of ordering". In: *Journal of Symbolic Computation* 16.4 (1993), pp. 329–344. ISSN: 0747-7171. DOI: *10.1006/jsco.1993.1051* (cit. on pp. 12, 72, 75, 80).

[Fau02] J.-C. Faugère. "A new efficient algorithm for computing Gröbner bases without reduction to zero (F_5)". In: *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation*. ACM, 2002, 75–83 (electronic). DOI: *10.1145/780506.780516* (cit. on pp. 8, 67, 68, 132).

[Fau10] J.-C. Faugère. "FGb: A Library for Computing Gröbner Bases". In: *Mathematical Software - ICMS 2010*. Ed. by K. Fukuda, J. Hoeven, M. Joswig, and N. Takayama. Vol. 6327. Lecture Notes in Computer Science. Kobe, Japan: Springer Berlin / Heidelberg, Sept. 2010, pp. 84–87. DOI: *10.1007/978-3-642-15582-6_17* (cit. on pp. 67, 87, 132).

[Fau99] J.-C. Faugère. "A new efficient algorithm for computing Gröbner bases (F_4)". In: *Journal of Pure and Applied Algebra* 139.1-3 (1999). Effective methods in algebraic geometry (Saint-Malo, 1998), pp. 61–88. ISSN: 0022-4049. DOI: *10.1016/S0022-4049(99)00005-5* (cit. on pp. 8, 67).

[FLP08] J.-C. Faugère, F. Levy-dit-Vehel, and L. Perret. "Cryptanalysis of Minrank". In: *Advances in Cryptology CRYPTO 2008*. Ed. by D. Wagner. Vol. 5157. Lecture Notes in Computer Science. Santa Barbara, CA, USA: Springer-Verlag, Aug. 2008, pp. 280–296. ISBN: 978-3-540-85173-8. DOI: *10.1007/978-3-540-85174-5_16* (cit. on p. 10).

[FM13] J.-C. Faugère and C. Mou. "Sparse FGLM algorithms". Anglais. Preprint available at *http://hal.inria.fr/hal-00807540*. 2013 (cit. on p. 80).

[Frö85] R. Fröberg. "An inequality for Hilbert series of graded algebras". eng. In: *Mathematica Scandinavica* 56 (1985), pp. 117–144. URL: *http://eudml.org/doc/166929* (cit. on p. 48).

[FS09] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009, pp. xiv+810. ISBN: 978-0-521-89806-5. DOI: *10.1017/CBO9780511801655*. URL: *http://dx.doi.org/10.1017/CBO9780511801655* (cit. on pp. 17, 29, 30).

[FS12] J.-C. Faugère and J. Svartz. "Solving Polynomial Systems Globally Invariant Under an Action of the Symmetric Group and Application to the Equilibria of N vortices in the Plane". In: *ISSAC '12: Proceedings of the 2012 international symposium on Symbolic and algebraic computation*. 2012, pp. 170–178. URL: *http://www-polsys.lip6.fr/~jcf/Papers/FS12.pdf* (cit. on p. 14).

[FSS11] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. "Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1, 1): algorithms and complexity". In: *Journal of Symbolic Computation* 46.4 (2011), pp. 406–437. ISSN: 0747-7171. DOI: *10.1016/j.jsc.2010.10.014* (cit. on p. 14).

[FSS12] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. "Critical Points and Gröbner Bases: the Unmixed Case". In: *Proceedings of the 2012 International Symposium on Symbolic and Algebraic Computation (ISSAC 2012)*. 2012, pp. 162–169 (cit. on p. 132).

[FSS13] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. "On the Complexity of the Generalized MinRank Problem". In: *Journal of Symbolic Computation* 55 (2013), pp. 30–58 (cit. on pp. 14, 132).

[FSV13] J.-C. Faugère, M. Safey El Din, and T. Verron. "On the complexity of computing Gröbner bases for quasi-homogeneous systems". In: *Proceedings of the 2013 International Symposium on Symbolic and Algebraic Computation*. ISSAC '13. Boston, USA: ACM, 2013 (cit. on pp. 83, 85).

[FSV15] J.-C. Faugère, M. Safey El Din, and T. Verron. "On the complexity of computing Gröbner bases for weighted homogeneous systems". In: (2015). To appear in Journal of Symbolic Computations. DOI: *10.1016/j.jsc.2015.12.001* (cit. on pp. 83, 85).

[Gau09] P. Gaudry. "Index calculus for abelian varieties of small dimension and the elliptic curve discrete logarithm problem". In: *Journal of Symbolic Computation* 44.12 (2009). Gröbner Bases in Cryptography, Coding Theory, and Algebraic Combinatorics, pp. 1690–1702. ISSN: 0747-7171. DOI: *10.1016/j.jsc.2008.08.005* (cit. on p. 120).

[Gio+91] A. Giovini et al. ""One sugar cube, please" or selection strategies in the Buchberger algorithm". In: *Proceedings of the 1991 international symposium on Symbolic and algebraic computation - ISSAC '91*. Vol. 91. 1991, pp. 49–54. ISBN: 0897914376. DOI: *10.1145/120694.120701* (cit. on p. 67).

[GKZ94] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants*. Boston, MA: Birkhäuser Boston, 1994. ISBN: 978-0-8176-4770-4, 978-0-8176-4771-1. DOI: *10.1007/978-0-8176-4771-1* (cit. on p. 32).

[GLS01] M. Giusti, G. Lecerf, and B. Salvy. "A Gröbner Free Alternative for Polynomial System Solving". In: *Journal of Complexity* 17.1 (2001), pp. 154–211. ISSN: 0885-064X. DOI: *10.1006/jcom.2000.0571* (cit. on p. 8).

[GM89] P. Gianni and T. Mora. "Algebraic solution of systems of polynomial equations using Groebner bases". In: *Applied algebra, algebraic algorithms and error-correcting codes (Menorca, 1987)*. Vol. 356. Lecture Notes in Comput. Sci. Springer, Berlin, 1989, pp. 247–257. DOI: *10.1007/3-540-51082-6_83*. URL: *http://dx.doi.org/10.1007/3-540-51082-6_83* (cit. on p. 64).

[GR09] E. Guerrini and A. Rimoldi. "FGLM-Like Decoding: from Fitzpatrick's Approach to Recent Developments". English. In: *Gröbner Bases, Coding, and Cryptography*. Ed. by M. Sala et al. Springer Berlin Heidelberg, 2009, pp. 197–218. ISBN: 978-3-540-93805-7. DOI: *10.1007/978-3-540-93806-4_12* (cit. on p. 87).

[Har77] R. Hartshorne. *Algebraic geometry*. Graduate Texts in Mathematics, No. 52. New York: Springer-Verlag, 1977, pp. xvi+496. ISBN: 0-387-90244-9 (cit. on pp. 38–40).

[HNS15a] D. Henrion, S. Naldi, and M. Safey El Din. "Exact algorithms for linear matrix inequalities". Submitted. Aug. 2015. URL: *https://hal.archives-ouvertes.fr/hal-01184320* (cit. on pp. 131, 132).

[HNS15b] D. Henrion, S. Naldi, and M. Safey El Din. "Real root finding for low rank linear matrices". Submitted. June 2015. URL: *https://hal.archives-ouvertes.fr/hal-01159210* (cit. on pp. 51, 131, 132, 139).

[HNS15c] D. Henrion, S. Naldi, and M. Safey El Din. "Real root finding for rank defects in linear Hankel matrices". In: *Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation*. Bath, United Kingdom, July 2015. URL: *https://hal.archives-ouvertes.fr/hal-01114378* (cit. on pp. 131, 132).

[HNS16] D. Henrion, S. Naldi, and M. Safey El Din. "Real root finding for determinants of linear matrices". In: *Journal of Symbolic Computation* 74 (2016), pp. 205–238. ISSN: 0747-7171. DOI: *http://dx.doi.org/10.1016/j.jsc.2015.06.010* (cit. on pp. 131, 132).

[HS12] H. Hong and M. Safey El Din. "Variant quantifier elimination". In: *Journal of Symbolic Computation* 47.7 (2012). International Symposium on Symbolic and Algebraic Computation (ISSAC 2009), pp. 883–901. ISSN: 0747-7171. DOI: *http://dx.doi.org/10.1016/j.jsc.2011.05.014* (cit. on p. 131).

[Lap+12] M. Lapert et al. "Exploring the Physical Limits of Saturation Contrast in Magnetic Resonance Imaging". In: *Scientific Reports* 2.589 (2012). DOI: *10.1038/srep00589*. URL: *http://dx.doi.org/10.1038/srep00589* (cit. on p. 18).

[Laz83] D. Lazard. "Gröbner-Bases, Gaussian elimination and resolution of systems of algebraic equations". In: *EUROCAL*. Ed. by J. A. van Hulzen. Vol. 162. Lecture Notes in Computer Science. Springer, 1983, pp. 146–156. ISBN: 3-540-12868-9. DOI: *10.1007/3-540-12868-9_99* (cit. on pp. 13, 67).

[LeGall14] F. Le Gall. "Powers of Tensors and Fast Matrix Multiplication". In: *CoRR* (2014). arXiv: *1401/7714* (cit. on p. 75).

[Leo09] D. A. Leonard. "A weighted module view of integral closures of affine domains of type I". In: *Advances in Mathematics of Communications* 3.1 (2009), pp. 1–11. ISSN: 1930-5346. DOI: *10.3934/amc.2009.3.1* (cit. on p. 87).

[LMX05] F. Lemaire, M. Moreno Maza, and Y. Xie. "The RegularChains library". In: *Maple conference*. Vol. 5. 2005, pp. 355–368 (cit. on pp. 8, 132).

[LR07] D. Lazard and F. Rouillier. "Solving parametric polynomial systems". In: *Journal of Symbolic Computation* 42.6 (2007), pp. 636–667 (cit. on pp. 19, 131, 135).

[Luc91] É. Lucas. *Théorie des nombres*. Vol. 1. Théorie des nombres. Gauthier-Villars et fils, 1891 (cit. on p. 93).

[Macaulay2] D. R. Grayson and M. E. Stillman. *Macaulay2, a software system for research in algebraic geometry*. Available at *http://www.math.uiuc.edu/Macaulay2/*. 2014 (cit. on pp. 67, 89).

[Magma] W. Bosma, J. Cannon, and C. Playoust. "The Magma algebra system. I. The user language". In: *Journal of Symbolic Computation* 24.3-4 (1997). Computational algebra and number theory (London, 1993), pp. 235–265. ISSN: 0747-7171. DOI: *10.1006/jsco.1996.0125* (cit. on pp. 67, 84, 87).

[Maple] Waterloo Maple (Maplesoft). *Computer Algebra System*. URL: *www.maplesoft.com/products/maple/* (cit. on p. 67).

[McC99] S. McCallum. "On projection in CAD-based quantifier elimination with equational constraint". In: *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation (Vancouver, BC)*. ACM, New York, 1999, 145–149 (electronic). DOI: *10.1145/309831.309892* (cit. on p. 131).

[MM82] E. W. Mayr and A. R. Meyer. "The complexity of the word problems for commutative semigroups and polynomial ideals". In: *Advances in Mathematics* 46.3 (Dec. 1982), pp. 305–329. ISSN: 00018708. DOI: *10.1016/0001-8708(82)90048-2* (cit. on p. 76).

[Mor03] G. Moreno-Socías. "Degrevlex Gröbner bases of generic complete intersections". In: *Journal of Pure and Applied Algebra* 180 (2003), pp. 263–283. URL: *http://www.sciencedirect.com/science/article/pii/S0022404902002979* (cit. on p. 48).

[Mor96] G. Moreno-Socías. *Revlex standard bases of generic complete intersections*. Technical report. 1996 (cit. on pp. 15, 33, 86, 96–98).

[Nal15] S. Naldi. "Exact algorithms for determinantal varieties and semidefinite programming". PhD thesis. Sept. 2015 (cit. on pp. 14, 44, 45).

[NZM91] I. M. Niven, H. S. Zuckerman, and H. L. Montgomery. *An introduction to the theory of numbers*. Wiley, 1991. ISBN: 9780471625469 (cit. on p. 93).

[Ott+15] J. C. Ottem, K. Ranestad, B. Sturmfels, and C. Vinzant. "Quartic spectrahedra". en. In: *Mathematical Programming* 151.2 (July 2015), pp. 585–612. ISSN: 0025-5610, 1436-4646. DOI: *10.1007/s10107-014-0844-3* (cit. on p. 10).

[Par10] K. Pardue. "Generic sequences of polynomials". In: *Journal of Algebra* 324.4 (2010), pp. 579–590. ISSN: 0021-8693. DOI: *10.1016/j.jalgebra.2010.04.018* (cit. on pp. 37, 47).

[Pon+62] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962, pp. viii+360 (cit. on p. 133).

[RRR91] L. Reid, L. G. Roberts, and M. Roitman. "On complete intersections and their Hilbert functions". In: *Canadian Mathematical Bulletin* 34.4 (1991), pp. 525–535. ISSN: 0008-4395. DOI: *10.4153/CMB-1991-083-9* (cit. on pp. 112, 113).

[RT15] J. I. Rodriguez and X. Tang. "Data-Discriminants of Likelihood Equations". In: *Proceedings of the 2015 ACM on International Symposium on Symbolic and Algebraic Computation*. ISSAC '15. Bath, United Kingdom: ACM, 2015, pp. 307–314. ISBN: 978-1-4503-3435-8. DOI: *10.1145/2755996.2756649* (cit. on p. 131).

[SF82] J. J. Sylvester and F. Franklin. "A Constructive Theory of Partitions, Arranged in Three Acts, an Interact and an Exodion". In: *Amer. J. Math.* 5.1-4 (1882), pp. 251–330. ISSN: 0002-9327. DOI: *10.2307/2369545*. URL: *http://dx.doi.org/10.2307/2369545* (cit. on p. 29).

[Singular] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann. *SINGULAR 3-1-6 — A computer algebra system for polynomial computations*. *http://www.singular.uni-kl.de*. 2012 (cit. on pp. 67, 119).

[Spa12] P.-J. Spaenlehauer. "Résolution de systèmes multi-homogènes et déterminantiels". PhD thesis. Oct. 2012. URL: *http://www.pjspaenlehauer.net/data/these%7B%5C_%7Dspaenlehauer.pdf* (cit. on pp. 16, 41, 42).

[Spa14] P.-J. Spaenlehauer. "On the complexity of computing critical points with Gröbner bases". In: *SIAM Journal on Optimization* 24.3 (2014), pp. 1382–1401 (cit. on p. 132).

[Str69] V. Strassen. "Gaussian elimination is not optimal". In: *Numerische Mathematik* 13.4 (Aug. 1969), pp. 354–356. ISSN: 0029-599X. DOI: *10.1007/BF02165411* (cit. on p. 75).

[Stu08] B. Sturmfels. *Algorithms in Invariant Theory (Texts and Monographs in Symbolic Computation)*. 2nd ed.; VII, 197 pp.; 5 figs. Springer Publishing Company, Incorporated, 2008. ISBN: 3211774165, 9783211774168 (cit. on pp. 87, 125).

[SW05] A. J. Sommese and C. W. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific Publishing Company, Incorporated, 2005. ISBN: 9789812567727 (cit. on p. 8).

[Tra96] C. Traverso. "Hilbert Functions and the Buchberger Algorithm". In: *Journal of Symbolic Computation* 22.4 (1996), pp. 355–376. ISSN: 0747-7171. DOI: *http://dx.doi.org/10.1006/jsco.1996.0056* (cit. on pp. 18, 67, 84, 124).

[YHX01] L. Yang, X. Hou, and B. Xia. "A complete algorithm for automated discovering of a class of inequality-type theorems". In: *Sci. China Ser. F* 44.1 (2001), pp. 33–49. ISSN: 1009-2757. DOI: *10.1007/BF02713938* (cit. on pp. 19, 131, 135).

Index