Change of ordering for regular chains in positive dimension

Xavier Dahan   Xin Jin   Marc Moreno Maza   Éric Schost

Abstract

We discuss changing the variable ordering for a regular chain in positive dimension. This quite general question has applications going from implicitization problems to the symbolic resolution of some systems of differential algebraic equations.

We propose a modular method, reducing the problem to dimension zero and using Newton-Hensel lifting techniques. The problems raised by the choice of the specialization points, the lack of the (crucial) information of what are the free and algebraic variables for the new ordering, and the efficiency regarding the other methods are discussed. Strong hypotheses (but not unusual) for the initial regular chain are required. Change of ordering in dimension zero is taken as a subroutine.

Introduction

Lexicographic orders on polynomial rings are useful tools. Even if computing Gröbner bases for such orders is difficult, these Gröbner bases are well suited to answer fast and easily to many problems. Lexicographic orders are also a key component to define *regular chains* (see Definition 0.2 and [3]), which are well established objects for polynomial system solving [16, 14, 15].

Suppose that we are given a regular chain as input, as well as a *target* order on the variables; we are interested in converting *symbolically* this input into a new regular chain with respect to the *target* order, while describing the same solutions. This is required by many applications, ranging from implicitization problems to invariant theory, as in the following example.

EXAMPLE. Consider the polynomials P in $\mathbb{Q}[X_1, X_2]$ such that $P(X_1, X_2) = P(-X_1, -X_2)$. Invariant theory tells us that any such polynomial can be written as a polynomial in X_1^2, X_2^2 (the *primary* invariants π_1 and π_2) and X_1X_2 (the *secondary* invariant σ); natural questions to ask are whether such a representation is unique, and how to perform the rewriting. This can be done by getting an expression of X_1 and X_2 in function of π_1 and π_2, hence by changing the order of the following system from $\pi_1 > \pi_2 > \sigma > X_1 > X_2$ to $X_2 > X_1 > \sigma > \pi_1 > \pi_2$. Given

$$\left|\begin{array}{lcl} \pi_1 & = & X_1^2 \\ \pi_2 & = & X_2^2 \\ \sigma & = & X_1X_2 \end{array}\right. \qquad \text{or} \qquad \left|\begin{array}{lcl} \pi_1 - X_1^2 & = & 0 \\ \pi_2 - X_2^2 & = & 0 \\ \sigma - X_1X_2 & = & 0 \end{array}\right. ,$$

we obtain

$$\left|\begin{array}{lcl} \sigma X_2 - \pi_1 X_1 & = & 0 \\ X_1^2 - \pi_1 & = & 0 \\ \sigma^2 - \pi_1\pi_2 & = & 0 \end{array}\right. \qquad \text{or} \qquad \left|\begin{array}{lcl} X_2 & = & \frac{\sigma}{\pi_1}X_1 \\ X_1^2 & = & \pi_1 \\ \sigma^2 & = & \pi_1\pi_2 \end{array}\right. .$$

In this form, we observe the relation $\sigma^2 = \pi_1\pi_2$ between our basic invariants, which establishes that the representation cannot be unique. Furthermore, the new form of the system can be used as a set of *rewriting rules*, so as to obtain a *canonical form* for any invariant polynomial.

Main results. To state our results, we will make the following assumption:

(**H**) *the input is a regular chain whose saturated ideal is prime.*

Without this assumption, one may need *several* regular chains to describe our output; observe that the example above, and more generally, implicitization problems, satisfy this assumption.

Theorem 0.1. *Let $\mathbb{K}$ be a perfect field, let $\mathbf{C}$ be a regular chain in $\mathbb{K}[\mathbf{X}] = \mathbb{K}[X_1, \ldots, X_n]$, satisfying assumption $\mathbf{H}$, and let $<$ be an order on the variables $\mathbf{X}$. Suppose that the polynomials in $\mathbf{C}$ can be evaluated in L operations, and that the saturated ideal of $\mathbf{C}$ has dimension r and degree δ. There exists an algorithm that computes a regular chain $\mathbf{C}'$ for the order $<$, that admits the same saturated ideal as $\mathbf{C}$, and with the following features.*

The complexity is polynomial in L, δ, n, and in the number $\binom{r+\delta}{\delta}$ of monomials in r variables of degree at most δ. The algorithm chooses $O(n)$ random values in $\mathbb{K}$. Let d be the maximum degree of the polynomials in $\mathbf{C}$, and suppose that all random choices are done in a finite set $\Gamma \subset \mathbb{K}$, with the uniform distribution. Then the probability of failure is at most $2n(3d^n + n^2)d^{2n}/|\Gamma|$.

Previous work. In this paper, we focus on the case of *positive* dimension. There already exist many algorithms to perform the change of order in this context, see for instance [7, 4, 19]. Further, for the implicitization problem, which is an important application of change of order, there exist many specialized algorithms, relying on some form of resultant formalism or homological algebra techniques, see for instance [6, 2, 8] and the numerous references therein.

Our specificity is the use of *modular methods*, reducing as much as possible the positive-dimensional problems to zero-dimensional ones, following the philosophy of [12]. To do so, we rely on a few well-identified subroutines, such as change of order in dimension zero, and *Newton-Hensel lifting* to go back to positive dimension. Hence, most of the implementation effort is transferred to these central subroutines. Accordingly, though we do not do it for lack of space, one can state the complexity of our algorithm in terms of the cost of these subroutines only.

Overview of the algorithm

Definitions and first properties. Let $\mathbb{K}$ be a perfect field, let $\mathbf{X}$ be a set of n variables. Given a total order $\prec$ on $\mathbf{X}$, every non-constant polynomial in $\mathbb{K}[\mathbf{X}]$ can be viewed as univariate in its greatest variable; then, its *initial* is its leading coefficient.

Definition 0.2. Let $\mathbf{C} = C_1, \ldots, C_s$ be in $\mathbb{K}[\mathbf{X}]$ with respective (pairwise distinct) main variables $X_{i_1} \prec \cdots \prec X_{i_s}$. For all $1 \leq i \leq s$ the *saturated ideal* of $C_1, \ldots, C_i$ is the ideal $\langle C_1, \ldots, C_i \rangle : (h_1 \cdots h_i)^\infty$ where h_i is the initial of C_i. Then, the set $\mathbf{C}$ is a *regular chain* if for all $2 \leq i \leq s$ the initial h_i is regular modulo the saturated ideal of $C_1, \ldots, C_{i-1}$.

The main variables of the polynomials in $\mathbf{C}$ are its *algebraic* variables; the other variables are *free*. For $y \in \overline{\mathbb{K}}^{n-s}$, the *specialization* of $\mathbf{C}$ at y is obtained by evaluating the free variables at y in $\mathbf{C}$. For a generic value of y, it is a zero-dimensional regular chain in $\mathbb{K}[X_{i_1}, \ldots, X_{i_s}]$.

Many concepts used below are relevant from *matroid theory*. A matroid [20] is a combinatorial structure that captures the notion of independence (generalizing linear independence in vector spaces), and studies its combinatorial properties; it thus relates to notions of linear and algebraic independence, but also independence in graph theory.

Definition 0.3. A *matroid* M over a finite set $\mathbf{X}$ is given by a non-empty family $B(M)$ of subsets of $\mathbf{X}$ with the same cardinality r and satisfying the *exchange property*: for all $e, f \in B(M)$, for every

$v \in e - f$ there exists $w \in f - e$ such that $e - v + w \in B(M)$ holds. The elements of $B(M)$ are called the *bases* of M and r is its *rank*. The family of the $\mathbf{X} - e$, for all $e \in B$, is the set of the bases of a matroid M^* called the *dual matroid* of M.

As an example used below, consider a matrix A over $\mathbb{K}$, and suppose that the columns of A are indexed by $\mathbf{X}$. Let $B(A)$ be the set of all $\mathbf{Z} \subseteq \mathbf{X}$ such that: (1) the columns of $\mathbf{Z}$ are linearly independent and (2) $\mathbf{Z}$ is maximal for inclusion. Then, the elements of $B(A)$ are the bases of a matroid over $\mathbf{X}$, which we call the *linear matroid* generated by A over $\mathbb{K}$.

Let now $V \subset \overline{\mathbb{K}}^n$ be an irreducible algebraic variety defined over $\mathbb{K}$, and let $\mathcal{P} \subset \mathbb{K}[\mathbf{X}]$ be its defining ideal. Let r be the dimension of V, with $0 < r < n$, and define $s = n - r$. To such a variety, one can associate the matroid of the so-called "maximal sets of free variables".

Definition 0.4. Let $B(\mathbf{X})$ be the family of all $\mathbf{Y} \subseteq \mathbf{X}$ such that $\mathcal{P} \cap \mathbb{K}[\mathbf{Y}]$ equals $\langle 0 \rangle$, and such that $\mathbf{Y}$ is maximal for inclusion. The family $B(\mathbf{X})$ is the collection of the bases of a matroid on $\mathbf{X}$ of rank r, denoted by $\mathcal{M}_{\mathrm{coord}}(V)$ (the *coordinate matroid of* V).

This result shows how bases of $\mathcal{M}_{\mathrm{coord}}(V)$ are related to the descriptions of V by regular chains.

Theorem 0.5. *Let $\mathbf{Y}$ be a subset of $\mathbf{X}$ of cardinal r. Then, $\mathbf{Y}$ is a basis of $\mathcal{M}_{\mathrm{coord}}(V)$ if and only if there exists a regular chain $\mathbf{C}$ in $\mathbb{K}[\mathbf{X}]$ having $\mathcal{P}$ as saturated ideal and $\mathbf{Y}$ as free variables.*

Overview of the algorithm. Suppose now that we are given a regular chain $\mathbf{C}_0$ having $\mathcal{P}$ as saturated ideal. Assuming that the variables of $\mathbf{X}$ are ordered by a target order $<$, we aim at computing a regular chain for the target order, that has $\mathcal{P}$ as saturated ideal.

Our algorithm relies on the "specialize and lift" paradigm, in the following form. If $\mathbf{C}$ is any regular chain having $\mathcal{P}$ as saturated ideal, given a specialization $\mathbf{c}$ of $\mathbf{C}$, one can reconstruct $\mathbf{C}$ itself by applying the Newton-Hensel operator [12, 13, 18] to $\mathbf{c}$ and $\mathbf{C}_0$. Thus, we will first aim at computing a *specialization* of the output regular chain, starting from a *specialization* of the input one. As intermediate steps, we will consider a sequence of regular chains $\mathbf{C}_0, \ldots, \mathbf{C}_\ell$, where $\mathbf{C}_\ell$ is our output. At step i, the regular chain $\mathbf{C}_{i+1}$ is obtained from the current one $\mathbf{C}_i$ by exchanging the roles of two suitably chosen variables: a formerly *algebraic variable* v_i becomes free, and conversely, a formerly *free variable* w_i becomes algebraic. As above, these regular chains will be handled only through specializations $\mathbf{c}_i$.

As mentioned above, knowing a specialization $\mathbf{c}_\ell$ of $\mathbf{C}_\ell$, the output $\mathbf{C}_\ell$ can be recovered by Newton-Hensel lifting. Until this last step, we will work only with varieties of dimension zero or one; this is the key to keeping the complexity under control.

Algorithmic details. The algorithm is divided into two steps. First, we determine the sequences of variables v_i, w_i. Let $\mathbf{Z}_1, \mathbf{Z}_2$ be two distinct bases of the dual matroid $\mathcal{M}^*_{\mathrm{coord}}(V)$ of $\mathcal{M}_{\mathrm{coord}}(V)$. We write $\mathbf{Z}_1 <_{\mathrm{lex}} \mathbf{Z}_2$ if the largest element of $(\mathbf{Z}_1 - \mathbf{Z}_2) \cup (\mathbf{Z}_2 - \mathbf{Z}_1)$ is in $\mathbf{Z}_2$.

Theorem 0.6. *Let $\mathbf{C}'$ be a regular chain for the ordering $<$, having $\mathcal{P}$ as saturated ideal in $\mathbb{K}[\mathbf{X}]$. The set of the main variables of $\mathbf{C}'$ is the maximum basis of $\mathcal{M}^*_{\mathrm{coord}}(V)$ for the ordering $<_{\mathrm{lex}}$.*

The matroid $\mathcal{M}^*_{\mathrm{coord}}(V)$ is not easily accessible *a priori*. The following theorem, mostly a translation of the Jacobian criterion, shows how to solve this problem by linearization. Let $\mathrm{Jac}(\mathbf{C}_0)$ be the Jacobian matrix of $\mathbf{C}_0$; then, the columns of $\mathrm{Jac}(\mathbf{C}_0)$ are indexed by the variables of $\mathbf{X}$.

Theorem 0.7. *There is an non-empty open subset $V' \subset V$ such that for any $x \in V'$, the linear matroid on $\mathbf{X}$, over the field $\overline{\mathbb{K}}$, defined by the specialization of $\mathrm{Jac}(\mathbf{C}_0)$ at x equals $\mathcal{M}^*_{\mathrm{coord}}(V)$.*

Let $\mathbf{Z}_0$ be the algebraic variables of $\mathbf{C}_0$, and let y be a random point in $\mathbb{K}^r$. Specializing $\mathbf{C}_0$ at y, we obtain a zero-dimensional regular chain $\mathbf{c}_0$ in $\mathbb{K}[\mathbf{Z}_0]$. Let next J_0 be the matrix obtained by specializing $\mathrm{Jac}(\mathbf{C}_0)$ at y. Using Theorem 0.7, and working modulo $\mathbf{c}_0$, we can use J_0 to test linear independence in $\mathcal{M}^*_{\mathrm{coord}}(V)$. Then, we will compute two sequences $v_1, \dots, v_\ell$ and $w_1, \dots, w_\ell$ satisfying the following requirements: the elements of the sequence $(\mathbf{Z}_i)_{i=0,\dots,\ell}$ defined by $\mathbf{Z}_i = \mathbf{Z}_{i-1} + v_i - w_i$ are in $\mathcal{M}^*_{\mathrm{coord}}(V)$, and satisfy $\mathbf{Z}_{i+1} >_{\mathrm{lex}} \mathbf{Z}_i$. The exchange property shows that $\mathbf{Z}_\ell$ is indeed the maximal element in $\mathcal{M}^*_{\mathrm{coord}}(V)$; Theorem 0.6 shows that $\mathbf{Z}_\ell$ are the algebraic variables of our output regular chain.

It remains to do the change of variables at the level of regular chains. At step i, knowing a specialization $\mathbf{c}_i$ of $\mathbf{C}_i$, we will compute a specialization $\mathbf{c}_{i+1}$ of $\mathbf{C}_{i+1}$ through the following steps. First, we change the order of the variables in $\mathbf{c}_i$, putting w_i as least variable. Since $\mathbf{c}_i$ has dimension zero, this can be done using several algorithms [4, 7, 11, 17]. Then, we lift the free variable v_i using the Newton-Hensel operator, obtaining a one-dimensional regular chain $\mathbf{c}'_{i+1}$. Finally, we specialize w_i at a random value in $\mathbf{c}'_{i+1}$, obtaining $\mathbf{c}_{i+1}$, back in dimension zero.

Complexity and error probability analysis. We give only a sketch of the complexity analysis, using the notation of Theorem 0.1. The first part (determining the v_i and w_i) consists in linear algebra operations modulo the zero-dimensional regular chain $\mathbf{c}_0$; its complexity is thus polynomial in n and δ. The second part (computing the $\mathbf{c}_i$) uses change of order in dimension zero, and Newton-Hensel lifting for a single variable at a time; its complexity is linear in the complexity of evaluation L of $\mathbf{C}_0$, and polynomial in n and δ.

The last part in the algorithm recovers the r-dimensional regular chain $\mathbf{C}_\ell$ from its specialization $\mathbf{c}_\ell$; here, the complexity becomes polynomial in the number of monomials that can appear in $\mathbf{C}_\ell$, inducing a polynomial dependence in $\binom{r+\delta}{\delta}$. Observe that using the *straight-line program* encoding for the output, we could make this cost polynomial in δ. However, actual implementations use the dense encoding, with the cost given above.

Random choices are made to find the initial specialization point y, the specialization values for the variables w_i, and in the stopping criterion for the Newton-Hensel operator. Keeping track of all possible degeneracies leads to the bound in Theorem 0.1, which allows us to get the result with a probability as close to 1 as wanted. The proof relies on bounds of the size of coefficients of triangular sets [10], and on Bézout's theorem. The value reported in Theorem 0.1 essentially grows like d^{3n}; we are currently investigating the question to reduce this dependence to d^{2n}.

Worked example. We can describe the main steps of our algorithm in the previous simple example; all specialization values will be 1. Here, $\mathbf{Z}_0$ is the set of algebraic variables $\{\pi_1, \pi_2, \sigma\}$. Through linear algebra, we determine the variables that we will have to change, $(v_1 = X_2, w_1 = \pi_2)$ and $(v_2 = X_1, w_2 = \pi_1)$, leading to the dual bases $\mathbf{Z}_1 = \{\pi_1, X_2, \sigma\}$ and $\mathbf{Z}_2 = \{X_1, X_2, \sigma\}$. Then, at the level of the regular chains themselves, the following operations take place.

Step 1.

$$\begin{array}{lcl}\pi_2 & = & 1\\ \pi_1 & = & 1\\ \sigma & = & 1\end{array} \xrightarrow[\text{Dimension zero}]{\text{Change of order}} \begin{array}{lcl}\sigma & = & 1\\ \pi_1 & = & 1\\ \pi_2 & = & 1\end{array} \xrightarrow{\text{Lift } X_2} \begin{array}{lcl}\sigma & = & X_2\\ \pi_1 & = & 1\\ \pi_2 & = & X_2^2\end{array}$$

variables π_2, π_1, σ with $X_1 \leftarrow 1, X_2 \leftarrow 1$

variables σ, π_1, π_2 with $X_1 \leftarrow 1, X_2 \leftarrow 1$

variables $\sigma, \pi_1, \pi_2, X_2$ with $X_1 \leftarrow 1$

Specialization $\pi_2 \leftarrow 1$

Step 2.

$$\begin{array}{lcl}\sigma & = & X_2\\ \pi_1 & = & 1\\ X_2^2 & = & 1\end{array} \xrightarrow[\text{Dimension zero}]{\text{Change of order}} \begin{array}{lcl}X_2 & = & \sigma\\ \sigma^2 & = & 1\\ \pi_1 & = & 1\end{array} \xrightarrow{\text{Lift } X_1} \begin{array}{lcl}X_2 & = & \frac{\sigma}{X_1}\\ \sigma^2 & = & X_1^2\\ \pi_1 & = & X_1^2\end{array}$$

variables σ, π_1, X_2 with $X_1 \leftarrow 1, \pi_2 \leftarrow 1$

variables X_2, σ, π_1 with $X_1 \leftarrow 1, \pi_2 \leftarrow 1$

variables X_2, σ, π_1, X_1 with $\pi_2 \leftarrow 1$

Specialization $\pi_1 \leftarrow 1$

Step 3.

$$\begin{array}{lcl}X_2 & = & \frac{\sigma}{X_1}\\ \sigma^2 & = & 1\\ X_1^2 & = & 1\end{array} \xrightarrow[\text{Dimension zero}]{\text{Change of order}} \begin{array}{lcl}X_2 & = & \sigma X_1\\ X_1^2 & = & 1\\ \sigma^2 & = & 1\end{array} \xrightarrow{\text{Lift } \pi_1, \pi_2} \begin{array}{lcl}X_2 & = & \frac{\sigma X_1}{\pi_1}\\ X_1^2 & = & \pi_1\\ \sigma^2 & = & \pi_1 \pi_2\end{array}$$

variables X_2, σ, X_1 with $\pi_1 \leftarrow 1, \pi_2 \leftarrow 1$

variables X_2, X_1, σ with $\pi_1 \leftarrow 1, \pi_2 \leftarrow 1$

variables $X_2, X_1, \sigma, \pi_1, \pi_2$

Conclusion. This work extends the scope of modular methods to the problem of change of ordering. For this first insight, the strong primality hypothesis was required, but our results may open ways to a less restrictive situation. A first implementation has been achieved in Maple's `RegularChains` library [15]; a better-tuned implementation and comparative results are works in progress.

References

[1] C. Alonso, J. Guiterrez and T. Recio. An implicitization algorithm with fewer variables. *Comput. Aided Geom. Des.*, 12:251-258, Elsevier, 1995.

[2] C. D'Andrea and A. Khetan. Implicitization of rational surfaces with toric varieties, 2003.

[3] P. Aubry, D. Lazard and M. Moreno Maza. On the Theories of Triangular Sets. *J. Symb. Comp.*, 28:105–124, 1999.

[4] F. Boulier, F. Lemaire and M. Moreno Maza. PARDI !, In ISSAC'01, pp. 38–47, ACM, 2001.

[5] F. Boulier, F. Lemaire and M. Moreno Maza. Well known theorems on triangular systems and the D5 principle, In *Transgressive Computing 2006*, University of Granada, Spain, 2006.

[6] L. Busé and M. Chardin. Implicitizing rational hypersurfaces using approximation complexes. *J. Symb. Comp.*, 40:1150-1168 (2005).

[7] S. Collart, M. Kalkbrener and D. Mall. Converting Bases with the Gröbner Walk. *J. Symb. Comp.*, 24(3-4):465–470, 1997.

[8] D. Cox. Curves, surfaces, and syzygies. Topics in algebraic geometry and geometric modeling, *Contemp. Math.* 334:131–150, Amer. Math. Soc., 2003.

[9] X. Dahan, M. Moreno Maza, É. Schost, W. Wu and Y. Xie. Lifting techniques for triangular decompositions. In ISSAC'05, pp. 108–115, ACM, 2005.

[10] X. Dahan and É. Schost. Sharp Estimates for Triangular Sets. In ISSAC'04, pp. 103–110, ACM, 2004.

[11] J.-C. Faugère, P. Gianni, D. Lazard and T. Mora. Efficient Computation of Zero-Dimensional Gröbner Bases by Change of Ordering *J. Symb. Comp.*, 16(4):329–344, 1993

[12] M. Giusti, J. Heintz, J. E. Morais, and L. M. Pardo. When polynomial equation systems can be solved fast? In *AAECC-11*, vol. 948 of *Lect. Not. in Comp. Sci.*, pp. 205–231. Springer, 1995.

[13] M. Giusti, G. Lecerf, and B. Salvy. A Gröbner free alternative for polynomial system solving. *J. Complexity*, 17(1):154–211, 2001.

[14] M. Kalkbrener. A generalized Euclidean algorithm for computing triangular representations of algebraic varieties. *J. Symb. Comp.*, 15:143–167, 1993.

[15] F. Lemaire, M. Moreno Maza and Y. Xie. The `RegularChains` library. In *Maple Conference 2005*, 355-368, I. Kotsireas Ed., 2005.

[16] M. Moreno Maza. On triangular decompositions of algebraic varieties. MEGA-2000 Conference, Bath, 2000.

[17] C. Pascal and É. Schost. Change of order for bivariate triangular sets. In ISSAC'06, pp. 277–284, ACM, 2006.

[18] É. Schost. Degree Bounds and Lifting Techniques for Triangular Sets. In ISSAC'02, pp. 238–245, ACM, 2002.

[19] Q.-N. Tran. Efficient Groebner walk conversion for implicitization of geometric objects. *Comput. Aided Geom. Des.*, 21(9):837-857, Elsevier, 2004.

[20] D. J. A. Welsh. *Matroid Theory.* Academic Press, 1976.

Xavier Dahan, LIX, École polytechnique, Palaiseau, France. `dahan@lix.polytechnique.fr`

Xin Jin, Computer Science Department, University of Western Ontario, London, Ontario, Canada. `xjin5@csd.uwo.ca`

Marc Moreno Maza, Computer Science Department, University of Western Ontario, London, Ontario, Canada. `moreno@csd.uwo.ca`

Éric Schost, Computer Science Department, University of Western Ontario, London, Ontario, Canada. `eschost@csd.uwo.ca`