JOURNAL OF COMPLEXITY **16**, 70–109 (2000)
doi:10.1006/jcom.1999.0529, available online at http://www.idealibrary.com on IDEAL®

Deformation Techniques for Efficient Polynomial Equation Solving[1]

Joos Heintz

Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. I, (1428) Buenos Aires, Argentina; and Departamento de Matemática, Estadística y Computación, Facultad de Ciencias, Universidad de Cantabria, Avda. de los Castros s/n, E-39071 Santander, Spain
E-mail: joos@mate.dm.uba.ar, heintz@matesco.unican.es

and

Teresa Krick, Susana Puddu, Juan Sabia, and Ariel Waissbein

Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pab. I, (1428) Buenos Aires, Argentina

Received November 15, 1998

Suppose we are given a parametric polynomial equation system encoded by an arithmetic circuit, which represents a generically flat and unramified family of zero-dimensional algebraic varieties. Let us also assume that there is given the *complete* description of the solution of a particular unramified parameter instance of our system. We show that it is possible to "move" the given particular solution along the parameter space in order to reconstruct—by means of an arithmetic circuit—the coordinates of the solutions of the system for an arbitrary parameter instance. The underlying algorithm is highly efficient, i.e., polynomial in the syntactic description of the input and the following geometric invariants: the number of solutions of a typical parameter instance and the degree of the polynomials occurring in the output. In fact, we prove a slightly more general result, which implies the previous statement by means of a well-known primitive element algorithm. We produce an efficient algorithmic description of the hypersurface obtained projecting polynomially the given generically flat family of varieties into a suitable affine space. © 2000 Academic Press

Key Words: polynomial equation system; arithmetic circuit; shape (or primitive element) lemma; Newton–Hensel iteration.

[1] Research was partially supported by the Argentinian and Spanish Grants UBACYT TW 80, PIP CONICET 4571, DGICYT PB96-0671-C02-02, ANPCyT 03-00000-01593.

0885-064X/00 $35.00

1. INTRODUCTION

Suppose we have a (nonparametric) equation system given by sparse polynomials and assume that this equation system defines a zero-dimensional variety. Replacing some of the coefficients in the given polynomials by indeterminates we obtain a parametric equation system. If we take care that this replacement of coefficients by indeterminates produces a generically flat and unramified family of zero-dimensional varieties with underlying finite morphism, we may *move* the indeterminate coefficients to a certain value (e.g., the value zero, introducing thus more sparsity in the moved equation system). If we are finally able to solve this new *simplified* equation system we will also be able to solve *efficiently* the *original* equation system. Typical examples of equation systems where this method can be applied easily are the so-called "Pham systems" (see [LV98, MP97]). By the method exposed in this paper we are able to solve such systems in time which is polynomial in their Bézout-number (without hiding an extra exponential factor of 2^n as in [MP97]). This example also shows that the good algebraic and geometric properties of Pham systems do not help much to produce low complexities for their solution which seems intrinsically high (namely exponential).

Since a long time there exist numerical and symbolic attempts to solve polynomial equation systems by means of deformation techniques based on a perturbation of the original equation system and subsequent path-following methods (see, e.g., [AS88, CG83, SS94, Can90, GH93, KP96, VH94] for the algebraically closed case and [GV88, HRS90, Ren92, BPR97] for the real case). The drawback of these methods is that they typically introduce spurious solutions which may be expensive to identify and to eliminate. In this sense our method is quite different because it requires and makes possible a more careful use of perturbations in such situations.

The original aim of this paper is a better understanding of the main algorithmic idea contained in the papers [GHH$^+$97, GHM$^+$98, GHMP97, Mor97, HMPS98, BGHM]. For this purpose it is necessary to isolate this main idea. In this way we will also obtain an explanation for the long list of assumptions on the input equation system which appears in all statements of the papers mentioned before. On the other hand these assumptions seem to be quite natural, in the sense that they occur in many mathematical and even practical applications of polynomial equation solving (e.g., in calibration problems of robots, see [Kov90]).

A secondary aim of our paper is the analysis of the relation of our symbolic method to solve polynomial equation systems with the numerical methods of [SS93a, SS93b, SS93c, SS96b, SS94] which are also based on ideas of deformation in combination with the classical Newton algorithm (see also [Cas97]).

In this contribution we limit ourselves to the problem of finding a way for uncoupling the variables of the original equation system. In view of the shape (or primitive element) lemma used in [KP96, GHM$^+$98], such an uncoupling can easily be extended to a complete solution of the given polynomial system.

An important outcome of this paper is the following "golden rule" (in the spirit of A. Schönhage [SGV94]):

> *Be careful in selecting the projections when dealing with zero-dimensional elimination problems. Avoid generic primitive elements*!

This paper aims to illustrate that by a careful choice of the direction of projection of a given zero-dimensional variety (or of a given generically flat and unramified family of zero-dimensional varieties) the degree of the output equation may be small and that this degree influences stronlgy the complexity of the algorithm.

This observation is the key point of our examples and might be decisive for future (hopefully successful) computer implementations (see [GHL$^+$97, TER97, Bru98]). Finally, let us mention the relation between our work and Groebner solving [Buc85, WB93, Mis93]): parametric equation solving is particularly intricate for rewriting based methods. The corresponding algorithms solve the parametric polynomial equation systems by means of comprehensive Groebner bases. Since the output polynomials of the algorithmic problems under considerations are typically dense and contain many (parameter) variables, any rewriting based method faces here a big complexity problem just for writing down the resulting output.

1.1. *Terminology*

Before stating our results precisely let us fix some terminology and notations.

If $f_1, ..., f_s$ are polynomials in n indeterminates with coefficients in the field of rational numbers $\mathbf{Q}$, we will denote by $(f_1, ..., f_s)$ the ideal spanned by them in $\mathbf{Q}[X_1, ..., X_n]$, and by $rad(f_1, ..., f_s)$ its radical ideal.

The algebraic varieties V we consider in the sequel will always be affine, their points having coordinates in $\mathbf{C}$, the field of complex numbers. These varieties will be defined by explicit polynomial equations having coefficients in $\mathbf{Q}$.

The coordinate ring of V is denoted by $\mathbf{Q}[V]$. The dimension of a variety V, denoted by $\dim(V)$, is defined in the usual way as the maximum of the Krull dimensions of its irreducible components. We call a variety equidimensional if all its components have the same dimension. Following

[Hei83] we define the degree deg (V) of the variety V as the sum of the degrees of all its irreducible components.

As usual, $\mathbf{A}^n$ denotes the n-dimensional affine space $\mathbf{C}^n$ equipped with its Zariski topology.

Finally, $\mathbf{Z}$ and $\mathbf{N}$ will denote the ring of rational integers and the set of natural numbers.

For the precise statement of our main result, it is necessary to explain what we mean by a "description of a geometric solution of a given zero-dimensional equation system (or variety)." In fact such a description is based on the main content of the *shape* (or *primitive element*) *lemma* for the description of the radical of a zero-dimensional ideal (see [Kro82; Zar95, Chap. 1; Nar66; GM89; Koe03; Mac16, Chap. II; CG83] for early forms of such shape lemmas). Our (implicit) use of such shape lemmas is algorithmic and follows [GH93, KP96, GHM$^+$98, GHH$^+$97, HMPS98, TER97].

Definition 1. Let $X := (X_1, ..., X_n)$ and Y be indeterminates over $\mathbf{Q}$. Let $f_1, ..., f_s \in \mathbf{Q}[X]$ be polynomials defining a zero-dimensional variety $V \subset \mathbf{A}^n$. A *description of a geometric solution of the system* $f_1 = 0, ..., f_s = 0$ (*or of the variety* V) is given by a (nonzero) $\mathbf{Q}$-linear form $U := \lambda_1 X_1 + \cdots + \lambda_n X_n$ and $n+1$ univariate polynomials $q, v_1, ..., v_n \in \mathbf{Q}[Y]$ such that the following conditions are satisfied:

(i) $V = \{(v_1(\eta), ..., v_n(\eta)); \eta \in \mathbf{C}, q(\eta) = 0\}$ and $\{\eta \in \mathbf{C}; q(\eta) = 0\} = \{U(\xi); \xi \in V\}$.

(ii) the polynomial q is monic of degree $\# V$ and the degrees of $v_1, ..., v_n$ are strictly less than $\# V$.

In such a description, we require U and $q, v_1, ..., v_n$ to be given by the arrays of their coefficients.

Let us here observe that the condition deg $q = \# V$ and condition (i) can be reduced to the usual Shape Lemma statement:

(iii) $\mathrm{rad}(f_1, ..., f_s) = (q(U), X_1 - v_1(U), ..., X_n - v_n(U))$.

1.2. *The Main Result*

We are now ready to state our main result. For the unfamiliarized reader, the notion of division-free arithmetic circuit will be made explicit in Subsection 1.3.

Theorem 2. *Let* $T := (T_1, ..., T_m)$, $X := (X_1, ..., X_n)$ *and* Y *be indeterminates over* $\mathbf{Q}$. *Let* $F_1, ..., F_n$ *and* $G \in \mathbf{Q}[T, X]$ *be polynomials given by a division-free arithmetic circuit* β *in* $\mathbf{Q}[T, X]$.

Suppose that the polynomials $F_1, ..., F_n$ *define an equidimensional subvariety* V *of* $\mathbf{A}^m \times \mathbf{A}^n = \mathbf{A}^{m+n}$ *of dimension* m, *and that the morphism of* *affine varieties $\pi\colon V \to \mathbf{A}^m$ induced by the canonical projection of $\mathbf{A}^m \times \mathbf{A}^n$ onto $\mathbf{A}^m$ is finite and surjective. Suppose furthermore that π is generically unramified (in the scheme-theoretical sense).*

Let $\tilde{\pi}\colon V \to \mathbf{A}^{m+1}$ be the (finite) morphism of affine varieties defined by $\tilde{\pi}(v) := (\pi(v), G(v))$ for $v \in V$ and let $P(T, Y) \in \mathbf{Q}[T, Y]$ be the minimal equation of the image variety $\tilde{\pi}(V)$ (interpreted as a closed affine subvariety of $\mathbf{A}^{m+1}$ of codimension 1).

Suppose that there is given a point $t = (t_1, ..., t_m) \in \mathbf{Q}^m$ satisfying the condition that its (finite) fiber $\pi^{-1}(t)$ is (scheme-theoretically) unramified and that G maps $\pi^{-1}(t)$ onto $\deg_Y P$ distinct points of $\mathbf{A}^1$. Assume finally that there is given a description of a geometric solution of the fiber $\pi^{-1}(t)$ (see Definition 1).

Then, if the length and the nonscalar depth of β are bounded by L and λ respectively, d is an upper bound for $\deg_X F_1, ..., \deg_X F_n$ and $D := \deg \pi$, there exists a division-free arithmetic circuit γ in $\mathbf{Q}[T]$ with the following properties:

(i) *γ computes the coefficients of the polynomial P with respect to the main variable Y.*

(ii) *γ has size $O(d^2 n^7 D^2 \deg_T^2 P \log(\deg_T P)\, L) + D^{O(1)} \deg_T^2 P \times \log(\deg_T P)$.*

(iii) *γ has nonscalar depth $O((\log n + \lambda) \log(\deg_T P) + \log D)$.*

Moreover, there exists a uniform family of arithmetic networks of asymptotically the same size and nonscalar depth than γ which produces γ from the following data:

— *the input arithmetic circuit β.*

— *the rational numbers which represent the coordinates of t and the given description of a geometric solution of the fiber $\pi^{-1}(t)$.*

We will say that the polynomials $F_1, ..., F_n, G$ describe the *general* instance of the elimination problem we are considering in Theorem 2. The solution of this general problem instance will be given by the polynomial P, and the output we are looking for is a division-free straight-line program γ in $\mathbf{Q}[T]$ which evaluates the coefficients of the polynomial P with respect to the variable Y. We call $T := (T_1, ..., T_m)$ the parameters and $X := (X_1, ..., X_n)$ the variables of our general problem instance. Each parameter point $\tau \in \mathbf{A}^m$ determines a *specific* problem instance which is given by the polynomials $F_1(\tau, X), ..., F_n(\tau, X), G(\tau, X)$ and which has a solution represented by the specialized polynomial $P(\tau, Y)$. Observe that the n-variate polynomial $P(\tau, G(\tau, X))$ vanishes on the zero-dimensional variety $\pi^{-1}(\tau)$. Thus, if we map by $G(\tau, X)$ the common zeroes of the polynomial equation system $F_1(\tau, X) = 0, ..., F_n(\tau, X) = 0$ into the affine space $\mathbf{A}^1$, the polynomial $P(\tau, Y)$

vanishes on each image point of this map and, for a generically chosen point $\tau \in \mathbf{A}^m$, exactly on them. Moreover, since the polynomial $P(T, Y)$ is monic in Y, we have $P(\tau, Y) \neq 0$ for any parameter point $\tau \in \mathbf{A}^m$. In this sense, the elimination problem we are considering (namely that of finding a straight-line program representation for the polynomial P starting with the given data $F_1, ..., F_n, G, t$ and $\pi^{-1}(t)$ is of *parametric* nature.

How do we proceed if our task is to find for an arbitrarily given parameter point $\tau \in \mathbf{Q}^m$ a description of a geometric solution of the zero-dimensional variety defined by the equations $F_1(\tau, X), ..., F_n(\tau, X)$? In this case we choose a generic linear form $U := \lambda_1 X_1 + \cdots + \lambda_n X_n$ of $\mathbf{Q}[X_1, ..., X_n]$ and compute solutions for the $2n$ specific problem instances $F_1(\tau, X), ..., F_n(\tau, X), X_i$ and $F_1(\tau, X), ..., F_n(\tau, X), U - \lambda_i X_i$, where $1 \leqslant i \leqslant n$. From these solution polynomials we compute now easily a description of the geometric solution of the polynomial equation system $F_1(\tau, X) = 0, ..., F_n(\tau, X) = 0$ applying the algorithms subjacent to [KP96, Lemma 26, Proposition 27, and Theorem 22].

In order to enlighten the long list of assumptions for our input equation system in Theorem 2, let us comment here some of its aspects:

(i) Under the assumption that V is equidimensional and that π is finite and surjective, the condition of π being generically unramified in the scheme-theoretical sense is equivalent to the requirement that the polynomials $F_1, ..., F_n$ span a radical ideal in $\mathbf{Q}(T)[X]$. Thus, under the assumptions of Theorem 2, if $F_1, ..., F_n$ form a regular sequence in $\mathbf{Q}[T, X]$, it turns out that the ideal $(F_1, ..., F_n)$ is a radical ideal of $\mathbf{Q}[T, X]$.

(ii) As the image of $\tilde{\pi}$ is an hypersurface of $\mathbf{A}^{m+1}$, its minimal polynomial $P(T, Y)$ is well defined. Since $\mathbf{Q}$ is a perfect field, the minimal polynomial $P(T, Y)$ is square free and since the morphism π is finite, $P(T, Y)$ is monic in Y. The polynomial $P(T, G)$ vanishes on the variety V and belongs to the ideal spanned by $F_1, ..., F_n$ in $\mathbf{Q}(T)[X]$. Moreover, if $F_1, ..., F_n$ is a regular sequence in $\mathbf{Q}[T, X]$, the polynomial $P(T, G)$ belongs to the ideal spanned by $F_1, ..., F_n$ in $\mathbf{Q}[T, X]$.

(iii) The assumption that the given parameter point $t = (t_1, ..., t_m) \in \mathbf{Q}^m$ has a scheme-theoretically unramified fiber $\pi^{-1}(t)$ means that the specialized ideal $(F_1(t, X), ..., F_n(t, X))$ is radical in $\mathbf{Q}[X]$. This, in turn, implies that the fiber $\pi^{-1}(t)$ contains only smooth points of the variety V, having regular local rings. Thus, the semilocal ring

$$\mathbf{Q}[T]_{(T_1 - t_1, ..., T_m - t_m)}[X]/(F_1, ..., F_n)$$

is Cohen–Macaulay and from the finiteness of π we deduce that

$$\# \pi^{-1}(t) = \dim_{\mathbf{Q}(T)} \mathbf{Q}(T)[X]/(F_1, ..., F_n)$$

holds (here $\#\pi^{-1}(t)$ denotes the cardinality of the set $\pi^{-1}(t)$ and $\dim_{\mathbf{Q}(T)} \mathbf{Q}(T)[X]/(F_1, ..., F_n)$ denotes the dimension of the $\mathbf{Q}(T)$-vector space $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$). In other words, $\#\pi^{-1}(t)$ equals the cardinality $D = \deg \pi$ of the generic fiber of π. This fact may also be circumscribed as the generic flatness of the morphism π.

Moreover the given description of a geometric solution of the fiber $\pi^{-1}(t)$ provides an explicit description of the multiplication tensor of the $\mathbf{Q}$-algebra $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$.

(iv) The assumption that G maps the fiber $\pi^{-1}(t)$ onto $\deg_Y P$ distinct points of $\mathbf{A}^1$ is equivalent to the requirement that the specialized polynomial $P(t, Y) := P(t_1, ..., t_m, Y) \in \mathbf{Q}[Y]$ is square-free, i.e., that the discriminant of $P(T, Y)$ as a polynomial in Y does not vanish on the point t.

(v) Finally observe that by our assumptions the squarefreeness of $P(T, Y)$ and the generical unramifiedness of π can be expressed as consistent Zariski open conditions in $\mathbf{A}^m$. Therefore, there always exists a point t satisfying our requirements.

Let us observe that in case $D \ll \deg V$ and $\deg P \ll \deg V.\deg G$ the algorithm underlying Theorem 2 is particularly efficient for the purpose of representing P by a "short" straight-line program.

In many (mathematical) applications, D and $\deg P$ are relatively small numbers whereas $\deg V$ tends to be close to the "global" Bézout-number $\bar{\delta} := \prod_{1 \leqslant i \leqslant n} \deg_{T,X} F_i$.

It may even occur that the "local" Bézout-number $\delta := \prod_{1 \leqslant i \leqslant n} deg_X F_i$ remains acceptable for computational purposes whereas $\bar{\delta}$ becomes exorbitant.

Let us recall that d was a given upper bound for $\deg_X F_1, ..., \deg_X F_n$. Let us suppose that there is also given an upper bound $\bar{d}$ for $\deg_{T,X} F_1, ..., \deg_{T,X} F_n$. With this notation one deduces easily from the Bézout Inequality in its simplest form (see, e.g, Hei83, Ful84, Vog84, Sch95, Som97, SS96a]) the following estimates:

$$D \leqslant \deg V, \qquad \delta \leqslant \bar{\delta},$$

$$\deg V \leqslant \bar{\delta} \leqslant \bar{d}^n, \qquad D \leqslant \delta \leqslant d^n,$$

$$\deg P \leqslant \deg V.\deg G \leqslant \bar{\delta}.\deg G \leqslant \bar{d}^n.\deg G,$$

$$\deg_Y P \leqslant D \leqslant \delta \leqslant d^n.$$

Let us also mention that, disregarding the aspect of nonscalar complexity, the estimates of Theorem 2 concerning the size of the straight-line program

γ and the arithmetic network producing γ may be refined to the more accurate bound

$$O(d^2 n^3 \deg_T^2 P \log(\deg_T P)\, L + D^2 \log D \log\log D \deg_T^2 P \log(\deg_T P)).$$

This means that sequential time, roughly quadratic in the degrees of the morphism π and the output polynomial P, and linear in the length of the input circuit β, suffices to represent the polynomial P by a straight-line program γ. The highlight of this complexity outcome is that the underlying algorithm can be executed in storage space which is almost *linear* in the geometric quantities n, D and $\deg_T P$ and the memory space necessary to evaluate the circuit β in time L (in all these considerations we are disregarding a logarithmic factor).

Since these accurate complexity results are not within the scope of this paper, we refer to [TER97], where suitable methods which allow to refine our algorithmic results in the indicated direction can be found.

Although in this paper we are not concerned with explicit bit complexity aspects, the nonscalar parallel and sequential time estimations of Theorem 2 allow to control suitably the height of the rational numbers occurring as intermediate and final results of our computations if the height of the parameters of the input circuit β and of the point t and its fiber $\pi^{-1}(t)$ are given. For a sufficiently generic input point t, we are even able to control the logarithmic height of the output polynomial P in this way (see [HMPS98]). On the other hand, the rational numbers representing the parameters of the input circuit β, the input point t and the given description of a geometric solution of the fiber $\pi^{-1}(t)$, may be given in their turn by a straight-line program α in $\mathbf{Q}$, which computes them from the input bits 0 and 1. Under these circumstances our complexity results do not change substantially. In particular, if the straight-line program α describing the data β, t and $\pi^{-1}(t)$ is again of length bounded by L, the asymptotic order of the sequential time of the algorithm underlying Theorem 2 remains unchanged.

Our paper contains also a refinement of the above mentioned result for the case that the Galois group of $\mathbf{C}$ over $\mathbf{Q}$ does not act transitively on the fiber $\pi^{-1}(t)$ (see Corollary 8 below). Finally, we illustrate the usefulness of our main result Theorem 2 and of our method by means of two classical computational examples.

1.3. *The Computational Model*

Our algorithms will be realized by *uniform* families of *arithmetic networks* (arithmetic Boolean circuits) over $\mathbf{Q}$. An arithmetic network is a pair $\Gamma = (G, Q)$, where G is a directed acyclic graph and Q is a labeling which assigns to each node of the graph G an arithmetic or Boolean operation.

A labeled node is also called a *gate* of Γ. A node of indegree zero may be labeled by a rational number or by an input variable. A node of indegree one or two may be labeled by an arithmetic operation $* \in \{+, -, \times, \div\}$ or a Boolean operation. Other nodes act as equality tests (decision gates) or are selector gates associated to an equality test. Some internal nodes of the graph G (among them all those of outdegree zero) are labeled as output gates of the arithmetic circuit Γ.

We shall always suppose that Γ is *division-free*. This means that when evaluating (running) Γ on a generic input (e.g., on its input variables), we will execute—if any—only divisions by nonzero rational numbers. This assumption implies that Γ computes only rational numbers and polynomials with rational coefficients in the input variables. If a node of Γ represents a multiplication of two polynomials of *positive* degree, we shall call *nonscalar* this multiplication and this node (gate). We count arithmetic or Boolean operations, equality tests and selections (performed by selector gates) only at unit cost. Thus, any arithmetic operation involving rational numbers or polynomials counts just as one unit, independently of the height of the rational numbers or degree of the polynomials we are processing. In this sense, we associate to a given arithmetic network $\Gamma = (G, Q)$ two complexity measures:

- *sequential time*, measured by the *size* of Γ (i.e., by the number of internal nodes G).
- *nonscalar parallel time*, measured by the *nonscalar depth* of Γ (i.e., by the longest oriented path in G, counting only nodes labeled by nonscalar multiplications).

Sometimes we will also refer to sequential time under storage space restrictions. In this case, time and space are measured by a pebble game on the graph G (see [Bor93]). If Γ is an arithmetic network without decision and selector gates (and, consequently, without Boolean operations), we call it an *arithmetic circuit* or *straight-line program*. Thus, our arithmetic circuits will always be division-free and will compute rational numbers and polynomials in the input variables having coefficients in $\mathbf{Q}$.

For the sake of simplicity, we restrict ourselves to the field of rational numbers $\mathbf{Q}$ as ground field, contained in the field of complex numbers $\mathbf{C}$. Nevertheless, our arguments can be easily extended to an *arbitrary* ground field of characteristic zero and any algebraically closed field containing it. With some slight modifications, our methods can be applied to any *infinite* (or finite, but sufficiently large) *perfect* ground field of *positive* characteristic and to its algebraic closure. In both cases, the complexity outcome remains essentially the same as before (see [GH93, GHM$^+$98, GHH$^+$97] for more details).

The (generally multivariate) polynomials over $\mathbf{Q}$ we deal with will be encoded in one of the following ways:

(i) in dense form, as arrays (vectors) of elements of $\mathbf{Q}$;

(ii) as arithmetic circuits (straight-line programs);

(iii) in *mixed representation*: in this case, the given polynomial is encoded in dense form with respect to a specific main variable whereas its coefficients with respect to this variable are encoded by an arithmetic circuit.

For precise definitions and elementary properties of the notions *arithmetic network* and *straight-line program* see [BCA97, vzG86, vzG93, KP96].

1.4. *Algorithmic Tools*

The algorithms we are going to design in this paper are based on the following three procedures:

(i) A particular symbolic adaptation of the classical Newton–Hensel iteration to the context of polynomial elimination and arithmetic circuits. This consists in fact in a refinement and a simplification of the main algorithmic idea of [GHH$^+$97, Lemma 30; GHM$^+$98; GHMP97; Mor97; HMPS98].

(ii) Linear algebra routines, like the well-parallelizable polynomial algorithm of Berkowitz for the computation of the characteristic polynomial of a square matrix over any domain [Ber84] (see also [vzG93, TER97]) and greatest common divisor (gcd) computations for multivariate polynomials given in mixed representation (see [Kal85, KP96]). This gcd computations may always be executed with sufficient efficiency by means of the algorithm of Berkowitz.

(iii) Suitable versions of Strassen's basic algorithm “Vermeidung von Divisionen” (see [Str73, KP96, GHH$^+$97]).

2. PROOFS

The procedure underlying Theorem 2 represents an algorithmic deformation of the given fiber $\pi^{-1}(t)$ along the parameter variety $\mathbf{A}^m$. Implicitly, this procedure makes use of the generic flatness of the morphism π.

The proof of Theorem 2 is based on the consideration of the multiplication tensor in the $\mathbf{Q}[T_1, ..., T_m]$-algebra $\mathbf{Q}[T_1, ..., T_m, X_1, ..., X_n]/rad(F_1, ..., F_n)$.

From now on, we will maintain the notations

$$T := (T_1, ..., T_m) \qquad \text{and} \qquad X := (X_1, ..., X_n).$$

Given a parameter point $t := (t_1, ..., t_m)$ belonging to $\mathbf{Q}^m$, we shall write $T - t := (T_1 - t_1, ..., T_m - t_m)$ and $\mathbf{Q}[\![T - t]\!] := \mathbf{Q}[\![T_1 - t_1, ..., T_m - t_m]\!]$ for the power series ring corresponding to $T - t$.

Observe that the following "algebraic" and "geometric" $\mathbf{Q}$-algebras are isomorphic:

$$\mathbf{Q}[T] \cong \mathbf{Q}[\mathbf{A}^m] \qquad \text{and} \qquad \mathbf{Q}[T, X]/rad(F_1, ..., F_n) \cong \mathbf{Q}[V].$$

Let us also observe that the generic unramifiedness and the finiteness of the epimorphism π imply that $\mathbf{Q}[T, X]/(F_1, ..., F_n) \cong \mathbf{Q}[V]$ holds if $F_1, ..., F_n$ is a regular sequence in $\mathbf{Q}[T, X]$.

By assumption, π is a finite morphism mapping the variety V onto the affine space $\mathbf{A}^m$. Therefore, $\mathbf{Q}[T, X]/rad(F_1, ..., F_n)$ is a finite $\mathbf{Q}[T]$-module containing $\mathbf{Q}[T]$ as a subalgebra.

Since $\mathbf{Q}[T]$ is integrally closed in its fraction field $\mathbf{Q}(T)$, the polynomial $P \in \mathbf{Q}[T, Y]$ we are looking for is the minimal equation of the residue class of G of the quotient ring $\mathbf{Q}[T, X]/rad(F_1, ..., F_n)$ over the subalgebra $\mathbf{Q}[T]$.

The way we compute the polynomial P is the following: first we approximate in a suitable way a matrix representation of the multiplication tensor of the (reduced) $\mathbf{Q}(T)$-algebra

$$\mathbf{Q}(T)[X]/(F_1, ..., F_n)$$

(the reduceness of this algebra follows from the assumption that π is generically unramified, finite and surjective).

For this purpose, we use a suitable $\mathbf{Q}(T)$-vector space basis of this $\mathbf{Q}(T)$-algebra. This basis will be determined by a suitably chosen primitive element $U = \lambda_1 X_1 + \cdots + \lambda_n X_n$ of $\mathbf{Q}[X]$. Then we compute a suitable approximation of the characteristic polynomial $\mathcal{X}_G(Y)$ of the $\mathbf{Q}(T)$-linear map induced by the multiplication by the residue class of G in $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$. From the finiteness of the morphism π we deduce that $\mathcal{X}_G$ belongs to the polynomial ring $\mathbf{Q}[T, Y]$. The given approximation of the characteristic polynomial $\mathcal{X}_G$ and the knowledge of the fiber $\pi^{-1}(t)$ (and hence of the polynomial $P(t, Y)$) allows us to compute an approximation to the polynomial P (see Lemma 7 below). From this approximation we compute an exact representation of the polynomial P by means of a procedure of the type "Vermeidung von Divisionen." The underlying approximation process is based on the previously mentioned symbolic adaptation of Newton–Hensel lifting in the power series ring $\mathbf{Q}[\![T - t]\!]$.

Let us now go further into the details of the proof of Theorem 2.

For this purpose, we explain briefly the ideas involving matrix computations in zero-dimensional algebras. We do this in the particular case of the unramified fiber $\pi^{-1}(t)$ whose coordinate ring is the reduced $\mathbf{Q}$-algebra

$$\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X)).$$

In order to simplify notations, let us write $\pi^{-1}(t) = \{t\} \times V_t$, where V_t is the zero-dimensional subvariety of $\mathbf{A}^n$ defined by

$$V_t := \{\xi \in \mathbf{C}^n; F_1(t, \xi) = 0, ..., F_n(t, \xi) = 0\}.$$

By assumption, we have at our disposal a description of a geometric solution of V_t, i.e., a primitive element $U = \lambda_1 X_1 + \cdots + \lambda_n X_n$ in $\mathbf{Q}[X]$ and polynomials $q, v_1, ..., v_n \in \mathbf{Q}[Y]$ such that the conditions (i) and (ii) of Definition 1 are satisfied. The polynomial q is monic and has degree $\deg q = \# V_t = \# \pi^{-1}(t) = D$ (recall that $\pi^{-1}(t)$ unramified implies $\# \pi^{-1}(t) = D$). Moreover, q is separable and the degrees of the polynomials $v_1, ..., v_n$ are strictly less than D.

Furthermore, since $(F_1(t, X), ..., F_n(t, X))$ is a radical ideal, we have

$$(F_1(t, X), ..., F_n(t, X)) = (q(U), X_1 - v_1(U), ..., X_n - v_n(U)). \tag{1}$$

This implies that

$$\mathbf{Q}[V_t] \cong \mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X)) \cong \mathbf{Q}[Y]/(q(Y))$$

holds. In particular $\mathbf{Q}[V_t]$ is a finite dimensional $\mathbf{Q}$-algebra of dimension D. Let us denote the residue class of the linear form U in $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$ by u. Note that we may also interpret u as the restriction $U|_{V_t}$ of the linear form U to the variety V_t.

The morphism $u: V_t \to \mathbf{A}^1$ separates the points of V_t and it is clear that $B := \{1, u, ..., u^{D-1}\}$ is a $\mathbf{Q}$-vector space basis of $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$.

Let us now write $V_t = \{\xi^{(\ell)}; 1 \leqslant \ell \leqslant D\}$, with $\xi^{(\ell)} = (\xi_1^{(\ell)}, ..., \xi_n^{(\ell)}) \in \mathbf{A}^n$, and $\pi^{-1}(t) = \{(t, \xi^{(\ell)}); 1 \leqslant \ell \leqslant D\}$. Since the coordinate function $u \in \mathbf{Q}[V_t]$ separates the points of the variety V_t, the D distinct roots of the univariate polynomial q (which are algebraic numbers belonging to $\mathbf{C}$) coincide with the image of V_t under the map u. Thus we have $\{y \in \mathbf{C};\ q(y) = 0\} = \{u(\xi); \xi \in V_t\}$. Moreover, from (1) we deduce that for $1 \leqslant \ell \leqslant D$ the identity

$$\xi^{(\ell)} = (v_1(u(\xi^{(\ell)})), ..., v_n(u(\xi^{(\ell)}))) \tag{2}$$

holds.

Considering the $\mathbf{Q}$-linear homothety η of $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$ induced by the multiplication by u, we see that the matrix M of η with respect to the basis B is the companion matrix of the polynomial q. The eigenvalues of this matrix are exactly the zeroes of the polynomial q, which are all distinct and identical to the D values $u(\xi^{(1)}), ..., u(\xi^{(D)})$. Therefore, there exists a $\mathbf{C}$-vector space basis B' of the $\mathbf{C}$-algebra $\mathbf{C}[X]/(F_1(t, X), ..., F_n(t, X))$ such that the matrix M' of the homothety η with respect to the basis B' has the diagonal form

$$M' = \begin{pmatrix} u(\xi^{(1)}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & u(\xi^{(D)}) \end{pmatrix}. \tag{3}$$

Let us now consider the $\mathbf{Q}[T]$-algebra

$$\mathbf{Q}[T, X]/(F_1(t, X), ..., F_n(t, X)) = \mathbf{Q}[T] \otimes \mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$$

which is clearly a free $\mathbf{Q}[T]$-module having the same basis B as the $\mathbf{Q}$-algebra $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$.

Observe also that B' is a basis of the free $\mathbf{C}[T]$-module $\mathbf{C}[T, X]/(F_1(t, X), ..., F_n(t, X))$.

For any polynomial H of $\mathbf{Q}[X]$ or $\mathbf{Q}[T, X]$, we may consider the homothety η_H induced by the multiplication by the residue class h of H in the corresponding algebra.

We shall denote the matrices of η_H with respect to the basis B and B' by M_H and M'_H respectively. These matrices have the same characteristic and minimal polynomials since they are similar.

Observe that the homotheties $\eta_{X_1}, ..., \eta_{X_n}$ (and therefore the matrices $M_{X_1}, ..., M_{X_n}$ and $M'_{X_1}, ..., M'_{X_n}$) commute. This implies that, for any polynomial H in $\mathbf{Q}[T, X]$

$$\eta_H = H(T\eta_{X_1}, ..., \eta_{X_n}), \qquad M_H = H(TM_{X_1}, ..., M_{X_n}), \qquad \text{and}$$
$$M'_H = H(TM'_{X_1}, ..., M'_{X_n}).$$

holds.

Let us now analyze the homotheties and matrices $\eta_{X_1}, ..., \eta_{X_n}$, $M_{X_1}, ..., M_{X_n}$ and $M'_{X_1}, ..., M'_{X_n}$.

From (1) we deduce the following identities between $\mathbf{Q}$-linear endomorphisms of $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$ and matrices over $\mathbf{Q}$ and $\mathbf{C}$:

$$\begin{aligned} \eta_{X_1} &= v_1(\eta), ..., \eta_{X_n} = v_n(\eta) \\ M_{X_1} &= v_1(M), ..., M_{X_n} = v_n(M) \\ M'_{X_1} &= v_1(M'), ..., M'_{X_n} = v_n(M'). \end{aligned} \tag{4}$$

Thus, for any polynomial H in $\mathbf{Q}[T, X]$, we obtain the following identities between $\mathbf{Q}[T]$-linear endomorphisms of $\mathbf{Q}[T, X]/(F_1(t, X), ..., F_n(t, X))$ and matrices over $\mathbf{Q}[T]$ and $\mathbf{C}[T]$:

$$\begin{aligned} \eta_H &= H(T, v_1(\eta), ..., v_n(\eta)), \\ M_H &= H(T, v_1(M), ..., v_n(M)) \\ M'_H &= H(T, v_1(M'), ..., v_n(M')). \end{aligned} \tag{5}$$

Since M' is a diagonal matrix, M'_H is a diagonal matrix too and we have

$$M'_H = \begin{pmatrix} H(T, v_1(u(\xi^{(1)})), ..., v_n(u(\xi^{(1)}))) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H(T, v_1(u(\xi^{(D)})), ..., v_n(u(\xi^{(D)}))) \end{pmatrix}.$$

From (2), we deduce the representation

$$M'_H = \begin{pmatrix} H(T, \xi^{(1)}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H(T, \xi^{(D)}) \end{pmatrix}.$$

In particular, for $G(t, X)$, we have

$$M'_{G(t, X)} = \begin{pmatrix} G(t, \xi^{(1)}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & G(t, \xi^{(D)}) \end{pmatrix}.$$

Therefore, $G(t, \xi^{(1)}), ..., G(t, \xi^{(D)})$ are the eigenvalues of the homothety $\eta_{G(t, X)}$ and hence of the matrix $M_{G(t, X)}$. Since these values (eliminating repetitions) are exactly the zeroes of the separable polynomial $P(t, Y)$ and the homothety $\eta_{G(t, X)}$ is diagonalizable, we deduce that $P(t, Y)$ is the minimal polynomial of $M_{G(t, X)}$.

Let us now take a little distance from our detailed analysis of matrix computations in the $\mathbf{Q}$-algebra $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$, which represents the coordinate ring of the zero-dimensional variety V_t and the fiber $\pi^{-1}(t)$. The outcome of our analysis was the conclusion that the specialized polynomial $P(t, Y)$ is the minimal polynomial of the homothety $\eta_{G(t, X)}$, whose representation with respect to the basis B is the matrix $M_{G(t, X)}$. Our arguments were based on two fundamental assumptions:

— the fiber $\pi^{-1}(t)$ is zero-dimensional

— the fiber $\pi^{-1}(t)$ is unramified.

The zero dimensionality and the unramifiedness of the fiber $\pi^{-1}(t)$ can be expressed algebraically by the fact that $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$ is a finite-dimensional, reduced $\mathbf{Q}$-algebra of dimension $D = \deg \pi$.

Hence, the same argumentation can be applied to the generic fiber of π, which is scheme-theoretically represented by the finite-dimensional reduced $\mathbf{Q}(T)$-algebra $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$ of dimension D. Therefore we conclude that the polynomial P is the minimal polynomial of the homothety η_G in $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$ induced by the multiplication by the residue class of G (observe that the finiteness and the surjectivity of the morphism π and the fact that $\mathbf{Q}[T]$ is integrally closed in its fraction field $\mathbf{Q}(T)$ implies that the minimal polynomial of η_G belongs, in fact, to the $\mathbf{Q}$-algebra $\mathbf{Q}[T, Y]$).

The question is now: How can we compute the minimal polynomial of the homothety η_G without determining first a matrix form of the multiplication tensor of the $\mathbf{Q}(T)$-algebra $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$?

This computation will be done by means of an approximation method based on the Newton–Hensel procedure.

Our first aim is to "lift" algorithmically each point $\xi = (\xi_1, ..., \xi_n)$ of $V_t = \pi^{-1}(t)$ to a n-tuple $R^{(\xi)} := (R_1^{(\xi)}, ..., R_n^{(\xi)})$ of power series of $\mathbf{C}[\![T - t]\!]$ such that the following two conditions hold:

- $F_1(T, R^{(\xi)}) = 0, ..., F_n(T, R^{(\xi)}) = 0$
- $R^{(\xi)}(t) := (R_1^{(\xi)}(t), ..., R_n^{(\xi)}(t)) = (\xi_1, ..., \xi_n) = \xi.$

We will show that, by our unramifiedness assumptions on π and $\pi^{-1}(t)$ and Lemma 3 below, such a lifting procedure always exists.

Thus, the D distinct points $R^{(\xi^{(1)})}, ..., R^{(\xi^{(D)})}$ of $\mathbf{C}[\![T - t]\!]^n$ have coordinates which are algebraic over $\mathbf{Q}(T)$ and they form a complete solution set of the zero-dimensional equation system

$$\{F_1(T, X) = 0, ..., F_n(T, X) = 0\}$$

over the ground field $\mathbf{Q}(T)$.

From our previous argumentation one deduces immediately that $P(T, Y)$ is the minimal polynomial of the diagonal matrix

$$\begin{pmatrix} G(T, R^{(\xi^{(1)})}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & G(T, R^{(\xi^{(D)})}) \end{pmatrix}.$$

This matrix represents the homothety η_G in a suitable $\mathbf{C}(T)$-vector space basis of $\mathbf{C}(T)[X]/(F_1, ..., F_n)$.

Let $\mathcal{M} := (T_1 - t_1, ..., T_m - t_m)$ be the maximal ideal of $\mathbf{C}[T]$ corresponding to the parameter point t.

Denote by $\mathbf{C}[T]_{\mathscr{M}}$ and $\mathbf{C}[T]_{\mathscr{M}}\,[X]/(F_1, ..., F_n) = (\mathbf{C}[T][X]/(F_1, ..., F_n))_{\mathscr{M}}$ the localization by $\mathscr{M}$ of the corresponding $\mathbf{C}[T]$-modules.

Since the $\mathbf{C}$-algebra $\mathbf{C}[X]/(F_1(t, X), ..., F_n(t, X))$ is reduced we easily see that the Jacobian matrix

$$\begin{pmatrix} \dfrac{\partial F_1(t, X)}{\partial X_1} & \cdots & \dfrac{\partial F_1(t, X)}{\partial X_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_n(t, X)}{\partial X_1} & \cdots & \dfrac{\partial F_n(t, X)}{\partial X_n} \end{pmatrix}$$

is unimodular over this algebra (this means that the residue class of its determinant is a unit of the algebra).

From this we deduce that the Jacobian matrix

$$DF(X) := DF(T, X) := \begin{pmatrix} \dfrac{\partial F_1(T, X)}{\partial X_1} & \cdots & \dfrac{\partial F_1(T, X)}{\partial X_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_n(T, X)}{\partial X_1} & \cdots & \dfrac{\partial F_n(T, X)}{\partial X_n} \end{pmatrix} \qquad (6)$$

is also unimodular over the localized $\mathbf{C}$-algebra $\mathbf{C}[T]_{\mathscr{M}}\,[X]/(F_1, ..., F_n)$ and that this algebra is, therefore, reduced.

This allows us to state the announced lifting procedure, which is in fact a variant of the Newton–Hensel Lemma (see, e.g., [Ive73, Proposition 7.2]). Although this result is well known, we will give a self-contained and constructive proof in order to give a mathematical illustration of the particular algorithmic techniques we are using in this paper.

LEMMA 3. *Let assumptions and notations be as before. Then, for any point* $\xi = (\xi_1, ..., \xi_n) \in V_t$, *there exists a unique n-tuple* $R^{(\xi)} = (R_1^{(\xi)}, ..., R_n^{(\xi)}) \in \mathbf{C}[\![T-t]\!]^n$ *of formal power series such that the following two conditions are satisfied*:

- $F_1(T, R^{(\xi)}) = 0, ..., F_n(T, R^{(\xi)}) = 0$
- $R^{(\xi)}(t) := (R_1^{(\xi)}(t), ..., R_n^{(\xi)}(t)) = \xi.$

Proof. Our arguments follow the lines of [Mat97]. They are based on the iterated application of the Newton–Hensel operator we are going to explain now.

Let $F(X):=(F_1(T,X),...,F_n(T,X))$ and let $DF(X)$ be defined as in (6). The Newton–Hensel operator associated to $F(X)$ is the n-tuple $N_F(X)$ of rational functions of $\mathbf{Q}(T,X)$ defined by

$$N_F(X)^t := \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} - DF(X)^{-1} \cdot \begin{pmatrix} F_1(T,X) \\ \vdots \\ F_n(T,X) \end{pmatrix} = X^t - DF(X)^{-1} \cdot F(X)^t,$$

where ^{t} denotes transposition.

The remark made above on the determinant $JF(X)$ of the Jacobian matrix $DF(X)$ implies that the residue classes of the entries of the n-tuple $N_F(X)$ are elements of the $\mathbf{Q}$-algebra $\mathbf{Q}[T]_{\mathscr{M}}[X]/(F_1,...,F_n)$.

Given $\xi \in V_t$ we define recursively for each $k \in \mathbf{Z}_{\geqslant 0}$ an n-tuple $R^{(k,\xi)} = (R_1^{(k,\xi)},...,R_n^{(k,\xi)})$ of elements of $\mathbf{C}[T]_{\mathscr{M}}$ as

$$\begin{cases} R^{(0,\xi)} := (R_1^{(0,\xi)},...,R_n^{(0,\xi)}) := \xi \\ R^{(k,\xi)} := (R_1^{(k,\xi)},...,R_n^{(k,\xi)}) = N_F(R^{(k-1,\xi)}). \end{cases}$$

The definition may also be rephrased as

$$R^{(k,\xi)} := N_F^k(\xi) \qquad \text{for any} \quad k \in \mathbf{Z}_{\geqslant 0}, \tag{7}$$

where N_F^k denotes the k-fold iterated application of the Newton–Hensel operator.

Let $\mathscr{M}_{\mathscr{M}}$ denote the extension ideal of $\mathscr{M}$ in $\mathbf{C}[T]_{\mathscr{M}}$.

In order to show that our sequence $(R^{(k,\xi)})_{k \in \mathbf{Z}_{\geqslant 0}}$ is well defined, we will prove recursively the following assertions:

(I) $F_i(T, R^{(k,\xi)}) \in \mathscr{M}_{\mathscr{M}}^{2^k}$ for any $1 \leqslant i \leqslant n$ and an $k \in \mathbf{Z}_{\geqslant 0}$.

(II) $JF(T, R^{(k,\xi)}) := \det(DF(T, R^{(k,\xi)})) \notin \mathscr{M}_{\mathscr{M}}$.

Proof of the Assertions (I) *and* (II). Assertions (I) and (II) will be proved by induction on k.

The case $k=0$ follows from the assumption that the fiber $\pi^{-1}(t)$ is unramified and that ξ belongs to $\pi^{-1}(t)$.

Now, suppose that both assertions hold for $k \geqslant 0$. Thus induction hypothesis (II) implies that $R^{(k+1,\xi)}$ is well defined.

Let $Y:=(Y_1,...,Y_n)$ be new indeterminates, and write $(X-Y)$ for the ideal generated by $X_1-Y_1,...,X_n-Y_n$ in $\mathbf{C}[T,X,Y]$.

Let $R^{(k+1,\xi)} - R^{(k,\xi)} := (R_1^{(k+1,\xi)} - R_1^{(k,\xi)},...,R_n^{(k+1,\xi)} - R_n^{(k,\xi)})$.

Considering the formal Taylor expansion of F and of its first partial derivatives in Y we obtain, for any $1 \leqslant i \leqslant n$,

$$F_i(T, X) \equiv F_i(T, Y) + \sum_{j=1}^{n} \frac{\partial F_i}{\partial X_j}(T, Y) \cdot (X_j - Y_j) \qquad \operatorname{mod}(X - Y)^2. \tag{8}$$

Replacing X by $R^{(k+1, \xi)}$ and Y by $R^{(k, \xi)}$ we deduce from (8) the following congruence relations in $\mathbf{C}[T]_{\mathscr{M}}$:

$$\begin{aligned} F_i(T, R^{(k+1, \xi)}) \equiv F_i(T, R^{(k, \xi)}) & \\ + \sum_{j=1}^{n} \frac{\partial F_i}{\partial X_j}(T, R^{(k, \xi)}) \cdot (R_j^{(k+1, \xi)} - R_j^{(k, \xi)}) & \\ \operatorname{mod}(R^{(k+1, \xi)} - R^{(k, \xi)})^2. & \end{aligned} \tag{9}$$

Recall that we have by definition $R^{(k+1, \xi)} = N_F(R^{(k, \xi)})$. This implies the identity

$$\begin{pmatrix} R_1^{(k+1, \xi)} - R_1^{(k, \xi)} \\ \vdots \\ R_n^{(k+1, \xi)} - R_n^{(k, \xi)} \end{pmatrix} = -DF(T, R^{(k, \xi)})^{-1} \begin{pmatrix} F_1(T, R^{(k, \xi)}) \\ \vdots \\ F_n(T, R^{(k, \xi)}) \end{pmatrix}. \tag{10}$$

Taking into account the induction hypothesis (I) we deduce from identity (10) that for any $1 \leqslant i \leqslant n$

$$R_i^{(k+1, \xi)} - R_i^{(k, \xi)} \in \mathscr{M}_{\mathscr{M}}^{2^k} \tag{11}$$

holds.

Now, let $DF_i(T, X) := (\partial F_i / \partial X_1, ..., \partial F_i / \partial X_n)$ be the ith row of the matrix $DF(T, X)$. Multiplying the identity (10) by $DF_i(T, R^{(k, \xi)})$ we conclude

$$\begin{aligned} & DF_i(T, R^{(k, \xi)})(R^{(k+1, \xi)} - R^{(k, \xi)})^t \\ & \quad = DF_i(T, R^{(k, \xi)})(-DF(T, R^{(k, \xi)})^{-1}) \begin{pmatrix} F_1(T, R^{(k, \xi)}) \\ \vdots \\ F_n(T, R^{(k, \xi)}) \end{pmatrix} \\ & \quad = -(0, ..., 1, ..., 0) \begin{pmatrix} F_1(T, R^{(k, \xi)}) \\ \vdots \\ F_n(T, R^{(k, \xi)} \end{pmatrix} \\ & \quad = -F_i(T, R^{(k, \xi)}), \end{aligned}$$

where the value 1 occurs in the ith entry of the n-tuple $(0, ..., 1, ..., 0)$.

Finally, plugging this identities in the Taylor expansion (9), we obtain the congruence relations

$$F_i(T, R^{(k+1,\xi)}) \equiv F_i(T, R^{(k,\xi)}) - F_i(T, R^{(k,\xi)}) \equiv 0 \qquad \operatorname{mod}(R^{(k+1,\xi)} - R^{(k,\xi)})^2.$$

In conclusion we have for any $1 \leqslant i \leqslant n$, the ideal membership relation

$$F_i(T, R^{(k+1,\xi)}) \in (R^{(k+1,\xi)} - R^{(k,\xi)})^2$$

in the ring $\mathbf{C}[T]_{\mathscr{M}}$. From (11) we deduce now that $F_i(T, R^{(k+1,\xi)}) \in \mathscr{M}_{\mathscr{M}}^{2^{k+1}}$ holds for any $1 \leqslant i \leqslant n$. This proves assertion (I) for $k+1$.

In order to prove assertion (II) for $k+1$, we use again a formal Taylor expansion, but this time that of the polynomial $JF(X) := \det(DF(X))$. In the same manner as in (8) we deduce the congruence relation

$$JF(T, X) \equiv JF(T, Y) + \sum_{j=1}^{n} \frac{\partial JF}{\partial X_j}(T, Y) \cdot (X_j - Y_j) \qquad \operatorname{mod}(X - Y)^2.$$

Replacing in this expression X by $R^{(k+1,\xi)}$ and Y by $R^{(k,\xi)}$ we obtain

$$JF(T, R^{(k+1,\xi)}) \equiv JF(T, R^{(k,\xi)}) + \sum_{j=1}^{n} \frac{\partial JF}{\partial X_j}(R^{(k,\xi)}) \cdot (R_j^{(k+1,\xi)} - R_j^{(k,\xi)}) \qquad \operatorname{mod}(R^{(k+1,\xi)} - R^{(k,\xi)})^2.$$

From (11) we deduce that, for any $1 \leqslant i \leqslant n$, the ideal membership relation $R_i^{(k+1,\xi)} - R_i^{(k,\xi)} \in \mathscr{M}_{\mathscr{M}}^{2^k} \subset \mathscr{M}_{\mathscr{M}}$ holds in $\mathbf{C}[T]_{\mathscr{M}}$. By induction hypothesis (II) we have $JF(T, R^{(k,\xi)}) \notin \mathscr{M}_{\mathscr{M}}$. Therefore we infer from (12) that $JF(T, R^{(k+1,\xi)})$ does not belong to the ideal $\mathscr{M}_{\mathscr{M}}$. This shows assertion (II) for $k+1$ and finishes the proof of both assertions.

Observing now that (11) is true for any $k \in \mathbf{Z}_{\geqslant 0}$, we infer that, for $1 \leqslant i \leqslant n$, the sequence of rational functions $(R_i^{(\xi,k)})_{k \in \mathbf{Z}_{\geqslant 0}}$ converges to a power series $R_i^{(\xi)}$ of $\mathbf{C}[\![T-t]\!]$.

Let us write $R^{(\xi)} := (R_1^{(\xi)}, ..., R_n^{(\xi)})$. From assertion (I) we deduce that for $1 \leqslant i \leqslant n$ and any $k \in \mathbf{Z}_{\geqslant 0}$, the rational function $F_i(T, R^{(k,\xi)})$ belongs to the ideal $\mathscr{M}_{\mathscr{M}}^{2^k}$ of $\mathbf{C}[T]_{\mathscr{M}}$. This means that the identity $F_i(T, R^{(\xi)}) = 0$ holds in $\mathbf{C}[\![T-t]\!]$ for any $1 \leqslant i \leqslant n$.

Taking into account that by definition $R^{(0,\xi)} = \xi$ holds we deduce from (11) the identity $R^{(\xi)}(t) = \xi$.

This finishes the proof of the existence of the power series $R_1^{(\xi)}, ..., R_n^{(\xi)}$ contained in the statement of Lemma 3.

In order to prove their uniqueness, suppose there is given another n-tuple $\bar{R}^{(\xi)} := (\overline{R_1}^{(\xi)}, ..., \overline{R_n}^{(\xi)})$ of power series of $\mathbf{C}[\![T-t]\!]$, such that $F_i(T, \bar{R}^{(\xi)}) = 0$ and $\bar{R}^{\xi}(t) = \xi$ holds for any $1 \leqslant i \leqslant n$.

Then, replacing Y by $R^{(\xi)}$ and X by $\bar{R}^{(\xi)}$ in (8), we deduce that in $\mathbf{C}[\![T-t]\!]$ the congruence relation

$$F_i(T, \bar{R}^{(\xi)}) \equiv F_i(T, R^{(\xi)}) + \sum_{j=1}^{n} \frac{\partial F_i}{\partial X_j}(T, R^{(\xi)}) \cdot (\bar{R}_j^{(\xi)} - R_j^{(\xi)})$$
$$\bmod (\bar{R}^{(\xi)} - R^{(\xi)})^2$$

holds for any $1 \leqslant i \leqslant n$.

This implies the congruence relation

$$DF(T, R^{(\xi)}) \cdot (R^{(\xi)} - \bar{R}^{(\xi)})^t \equiv 0 \qquad \bmod (R^{(\xi)} - \bar{R}^{(\xi)})^2$$

in $\mathbf{C}[\![T-t]\!]$.

By assumption, for any $1 \leqslant i \leqslant n$ both series $R_i^{(\xi)}$ and $\bar{R}_i^{(\xi)}$ have the same constant term, namely $R_i^{(\xi)}(t) = \bar{R}_i^{(\xi)}(t) = \xi_i$. Moreover, by assertion (II), for any $k \in \mathbf{Z}_{\geqslant 0}$, the rational function $JF(T, R^{(k, \xi)})$ does not belong to the ideal $\mathscr{M}_{\mathscr{M}}$ of $\mathbf{C}[T]_{\mathscr{M}}$. This implies that $DF(T, R^{(\xi)})$ is a unimodular matrix with entries in $\mathbf{C}[\![T-t]\!]$. Therefore we conclude that, for any $1 \leqslant i \leqslant n$, the difference of series $R_i^{(\xi)} - \overline{R_i}^{(\xi)}$ must belong to the ideal $(T-t)^2$ of $\mathbf{C}[\![T-t]\!]$ (here $(T-t) = (T_1 - t_1, ..., T_m - t_m)$ denotes the maximal ideal of the local ring $\mathbf{C}[\![T-t]\!]$). Repeating inductively this argument, we deduce that $R_i^{(\xi)} - \overline{R_i}^{(\xi)}$ belongs to the ideal $(T-t)^{2^k}$ for any $k \in \mathbf{Z}_{\geqslant 0}$. This implies that $R_i^{(\xi)} = \overline{R_i}^{(\xi)}$ holds. ▮

Our main algorithmic tool will be the Newton–Hensel operator, but it is evident that we cannot use this operator infinitely often in a finite step procedure. In this sense, we have to fix a precision of approximation which limits the number of iterated applications of the operator. This leads to the following notion of approximation:

Definition 4. Let $\mathbf{K}$ be the field $\mathbf{Q}$ or $\mathbf{C}$. Let $t \in \mathbf{Q}^m$ be a parameter point and let $\Phi, \tilde{\Phi} \in \mathbf{K}[\![T-t]\!]$ be formal power series. Let $(T-t) := (T_1 - t_1, ..., T_m - t_m)$ be the maximal ideal of the local ring $\mathbf{K}[\![T-t]\!]$.

For $s \in \mathbf{N}$ we say that *$\tilde{\Phi}$ approximates Φ with precision s* in $\mathbf{K}[\![T-t]\!]$ if

$$\Phi \equiv \tilde{\Phi} \qquad \bmod (T-t)^s \qquad \textit{holds.}$$

If $Q, \tilde{Q} \in \mathbf{K}[\![T-t]\!][Y]$ are polynomials in a single variable Y (of the same formal degree) with coefficients being formal power series, we will say that

$\tilde{Q}$ *approximates* Q *with precision* s if each coefficient of $\tilde{Q}$ approximates the corresponding coefficient of Q with precision s.

Let us fix for the moment a nonnegative integer k and let us consider the k-fold iterated Newton–Hensel operator $N_F^k(X)$ introduced in the proof of Lemma 3 (recall that we have $N_F^0(X)^t = X^t$, $N_F^1(X)^t = X^t - DF^{-1} \cdot F^t$, etc).

Observe that there exist polynomials $g_1^{(k)}, ..., g_n^{(k)}$ and $h^{(k)}$ of $\mathbf{Q}[T, X]$ such that

$$N_F^k(X) = \left(\frac{g_1^{(k)}}{h^{(k)}}, ..., \frac{g_n^{(k)}}{h^{(k)}}\right) \tag{13}$$

holds and such that the residue class of $h^{(k)}$ is a unit of the reduced $\mathbf{Q}$-algebra $\mathbf{Q}[T]_{\mathscr{M}}[X]/(F_1, ..., F_n)$.

In the sequel we shall need the following complexity result (see [GHH$^+$97, Lemma 30]):

Lemma 5. *Let β be a division-free arithmetic circuit of size L and nonscalar depth λ computing $F_1, ..., F_n$ in $\mathbf{Q}[T, X]$, and let d be an upper bound for $\deg_X F_i$ $(1 \leqslant i \leqslant n)$.*

Suppose that the Jacobian matrix $DF(T, X)$ has a nonzero determinant. Then, there exists a division-free arithmetic circuit β_* in $\mathbf{Q}[T, X]$ of size $O(kd^2n^7L)$ and nonscalar depth $O(k(\log n + \lambda))$ which evaluates suitable polynomials $g_1^{(k)}, ..., g_n^{(k)}$ and $h^{(k)}$ of $\mathbf{Q}[T, X]$ such that the residue class of $h^{(k)}$ is a unit of the reduced $\mathbf{Q}$-algebra $\mathbf{Q}[T]_{\mathscr{M}}[X]/(F_1, ..., F_n)$ and such that

$$N_F^k(X) = \left(\frac{g_1^{(k)}}{h^{(k)}}, ..., \frac{g_n^{(k)}}{h^{(k)}}\right)$$

holds.

This arithmetic circuit can be produced by a uniform arithmetic network of asymptotically the same size and nonscalar depth as the output circuit β_*.

The index k being fixed, we may write

$$g_1 := g_1^{(k)}, ..., g_n := g_n^{(k)} \qquad \text{and} \qquad h := h^{(k)}.$$

Since the residue class of the polynomial h is a unit of $\mathbf{Q}[T]_{\mathscr{M}}[X]/(F_1, ..., F_n)$, the homothety $\eta_{h(t, X)}$ is a regular $\mathbf{Q}$-linear endomorphism of

$$\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X)).$$

Therefore, from (5) we deduce that the matrix $M_{h(t,X)} = h(t, v_1(M), ..., v_n(M)) \in \mathbf{Q}^{D \times D}$ is regular and that $h(T, v_1(M), ..., v_n(M))$ is a unimodular $D \times D$ matrix of $\mathbf{Q}[T]_{\mathcal{M}}^{D \times D}$.

Finally we conclude that the matrices

$$\begin{aligned} N_i &:= \frac{g_i}{h}(T, v_1(M), ..., v_n(M)) \\ &:= h(T, v_1(M), ..., v_n(M))^{-1} \, g_i(T, v_1(M), ..., v_n(M)) \end{aligned}$$

are well defined in $\mathbf{Q}[T]_{\mathcal{M}}^{D \times D}$ for $1 \leqslant i \leqslant n$.

We are now ready to state our first approximation result, which is a variant of [GHH$^+$97, Lemma 30]:

Lemma 6. *Let notations and assumptions be as before.*

Let η_G be the homothety induced by G in the $\mathbf{Q}(T)$-algebra $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$, and denote by $\mathcal{X}_G \in \mathbf{Q}[T, Y]$ its characteristic polynomial.

Write $\tilde{M}_G := G(T, N_1, ..., N_n) \in \mathbf{Q}[T]_{\mathcal{M}}^{D \times D}$ for the matrix obtained by substituting in G the variables $X_1, ..., X_n$ by the matrices $N_1, ..., N_n$, and denote by $\tilde{\mathcal{X}}_G \in \mathbf{Q}[T]_{\mathcal{M}}[Y]$ the characteristic polynomial of $\tilde{M}_G$

Then, interpreting $\mathcal{X}_G$ and $\tilde{\mathcal{X}}_G$ as elements of $\mathbf{Q}[\![T-t]\!][Y]$, we see that $\tilde{\mathcal{X}}_G$ approximates $\mathcal{X}_G$ with precision 2^k in $\mathbf{Q}[\![T-t]\!][Y]$.

Intuitively speaking, the previous lemma states that the coefficients of the "approximate" characteristic polynomial $\tilde{\mathcal{X}}_G$, obtained by means of the k-fold iteration of the Newton–Hensel operator N_F, approximate the coefficients of the "exact" characteristic polynomial $\mathcal{X}_G$ with precision 2^k in $\mathbf{Q}[\![T-t]\!]$.

Proof. For $1 \leqslant \ell \leqslant D$, consider $R^{(\xi^{(\ell)})} := (R^{(\xi^{(\ell)})}, ..., R^{(\xi^{(\ell)})}) \in \mathbf{C}[\![T-t]\!]^n$, the n-tuple of formal power series defined in Lemma 3.

Let $\overline{\mathbf{Q}(T)}$ be an algebraic closure of the function field $\mathbf{Q}(T)$ and consider the homothety η_G induced by the polynomial G in the $\mathbf{Q}(T)$-algebra $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$ and in the $\overline{\mathbf{Q}(T)}$-algebra $\overline{\mathbf{Q}(T)}\,[X]/(F_1, ..., F_n)$. With respect to a suitable $\overline{\mathbf{Q}(T)}$ vector space basis of the latter algebra, the homothety η_G can be represented by a diagonal matrix of the form

$$\begin{pmatrix} G(T, R^{(\xi^{(1)})}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & G(T, R^{(\xi^{(D)})}) \end{pmatrix}.$$

Therefore, we have $\mathcal{X}_G = \prod_{\ell=1}^{D} (Y - G(T, R^{(\xi^{(\ell)})}))$.

Now, recall from (3) that there exists an invertible matrix $C \in \mathbf{C}^{D \times D}$ such that

$$CMC^{-1} = M' = \begin{pmatrix} u(\xi^{(1)}) & & \\ & \ddots & \\ & & u(\xi^{(D)}) \end{pmatrix}$$

holds.

Let $1 \leqslant i \leqslant n$. One deduces easily the matrix identities

$$\begin{aligned} CN_iC^{-1} &= C\frac{g_i}{h}(T, v_1(M), ..., v_n(M))\, C^{-1} \\ &= \frac{g_i}{h}(T, C, v_1(M)\, C^{-1}, ..., Cv_n(M)\, C^{-1}) \\ &= \frac{g_i}{h}(T, v_1(CMC^{-1}), ..., v_n(CMC^{-1})). \end{aligned}$$

Therefore, by means of (2), (7), and (13) we deduce that

$$\begin{aligned} &CN_iC^{-1} \\ &= \begin{pmatrix} \frac{g_i}{h}(T, v_1(u(\xi^{(1)})), ..., v_n(u(\xi^{(1)}))) & & \\ & \ddots & \\ & & \frac{g_i}{h}(T, v_1(u(\xi^{(D)})), ..., v_n(u(\xi^{(D)}))) \end{pmatrix} \\ &= \begin{pmatrix} \frac{g_i}{h}(T, \xi^{(1)}) & & \\ & \ddots & \\ & & \frac{g_i}{h}(T, \xi^{(D)}) \end{pmatrix} \\ &= \begin{pmatrix} R_i^{(k, \xi^{(1)})} & & \\ & \ddots & \\ & & R_i^{(k, \xi^{(D)})} \end{pmatrix} \end{aligned}$$

holds.

In this way we obtain

$$C\tilde{M}_G C^{-1} = CG(T, N_1, ..., N_n)\, C^{-1} = G(T, CN_1C^{-1}, ..., CN_nC^{-1}),$$

which means

$$C\tilde{M}_G C^{-1} = \begin{pmatrix} G(T, R^{(k, \xi^{(1)})}) & & \\ & \ddots & \\ & & G(T, R^{(k, \xi^{(D)})}) \end{pmatrix}$$

From the similarity of the matrices $\tilde{M}_G$ and $C\tilde{M}_G C^{-1}$, we deduce that

$$\tilde{\mathcal{X}}_G(Y) = \prod_{1 \leqslant \ell \leqslant D} (Y - G(T, R^{(k, \xi^{(\ell)})}))$$

holds.

Now, from Lemma 3 and statement (11) of its proof it follows immediately that $\tilde{\mathcal{X}}_G$ approximates $\mathcal{X}_G$ with precision 2^k in $\mathbf{Q}[\![T-t]\!][Y]$. ▮

The next result will give us a way to compute *efficiently* the minimal polynomial $P(T, Y)$ of the homothety induced by G in the $\mathbf{Q}(T)$-algebra $\mathbf{Q}(T)[X]/(F_1, ..., F_n)$, from the knowledge of the approximating characteristic polynomial $\tilde{\mathcal{X}}_G$ and the minimal polynomial $P(t, Y)$ of the homothety $\eta_{G(t, X)}$ of $\mathbf{Q}[X]/(F_1(t, X), ..., F_n(t, X))$.

Lemma 7. *Let $Q(T, Y)$ and $\tilde{Q}(T, Y) \in \mathbf{Q}[T][Y]$ be polynomials, monic in the (single) variable Y satisfying the condition* $\deg_Y Q = \deg_Y \tilde{Q}$.

Write $R := (Q/\mathrm{gcd}_Y(Q, \partial Q/\partial Y)) \in \mathbf{Q}[T][Y]$.

Let $t \in \mathbf{Q}^m$ and assume that the following holds:

- $\deg_Y \mathrm{gcd}_Y(Q, \partial Q/\partial Y) = \deg \mathrm{gcd}(Q(t, Y), (\partial/\partial Y)\, Q(t, Y))$.
- *$\tilde{Q}$ approximates Q with positive precision s in $\mathbf{Q}[\![T-t]\!][Y]$ (thus we have in particular $\tilde{Q}(t, Y) = Q(t, Y)$).*

Suppose that the coefficients of $\tilde{Q} \in \mathbf{Q}[T][Y]$ are given by a division-free arithmetic circuit $\tilde{\beta}$ in $\mathbf{Q}[T]$ of size $\tilde{L}$ and nonscalar depth $\tilde{\lambda}$.

Then there exists a division-free arithmetic circuit $\tilde{\gamma}$ in $\mathbf{Q}[T]$ of size $\tilde{L} + (\deg_Y \tilde{Q})^{O(1)} \log s$ and nonscalar depth $\tilde{\lambda} + O(\log \deg_Y \tilde{Q} + \log s)$ which computes the coefficients in $\mathbf{Q}[T]$ of a polynomial $\tilde{R} \in \mathbf{Q}[T][Y]$ satisfying:

- $\deg_Y \tilde{R} = \deg_Y R$
- *$\tilde{R}$ approximates R with precision s in $\mathbf{Q}[\![T-t]\!][Y]$.*

The circuit $\tilde{\gamma}$ can be produced from the input circuit $\tilde{\beta}$ and from the coefficient representation of $Q(t, Y) = \tilde{Q}(t, Y)$ by a uniform arithmetic network of asymptotically the same size and nonscalar depth as $\tilde{\gamma}$.

Proof. All the polynomials occurring in this proof are considered as polynomials in the variable Y.

The idea of the proof consists in the observation that the polynomial $R(T, Y)$ can be represented by means of minors of a submatrix of the Sylvester matrix of Q and $\partial Q/\partial Y$. This submatrix can be made explicit analyzing the Sylvester matrix of $Q(t, Y) = \tilde{Q}(t, Y)$ and $\frac{\partial Q}{\partial Y}(t, Y) = \frac{\partial Q}{\partial Y}(t, Y)$. Then mimicking this representation of $R(T, Y)$ by means of the corresponding minors of the Sylvester matrix of $\tilde{Q}$ and $\partial\tilde{Q}/\partial Y$ we obtain the required approximation polynomial $\tilde{R}(T, Y)$.

Let us write $Q' := \partial Q/\partial Y$, $\gcd(Q, Q') := \gcd_Y(Q, Q')$, $d := \deg_Y Q(T, Y)$ and $r := \deg_Y R(T, Y)$.

Since by assumption Q is monic in Y and $\deg_Y \gcd(Q, Q') = \deg \gcd(Q(t, Y), Q'(t, Y))$ holds, we conclude $\gcd(Q(t, Y), Q'(t, Y)) = \gcd(Q, Q')(t, Y)$.

Therefore we have $R(t, Y) = (Q(t, Y)/\gcd(Q(t, Y), Q'(t, Y)))$, and $\deg R(t, Y) = r$.

Consider now the $\mathbf{Q}(T)$-linear mapping φ from $\mathbf{Q}(T)^{r+1} \oplus \mathbf{Q}(T)^r$ to $\mathbf{Q}(T)^{d+r}$ given by the polynomial expression

$$A(T, Y)\, Q'(T, Y) + B(T, Y)\, Q(T, Y)$$

and the degree restrictions $\deg_Y A \leqslant r$ and $\deg_Y B \leqslant r-1$. This $\mathbf{Q}(T)$-linear mapping has a one-dimensional kernel which can be described by the polynomial expressions

$$A(T, Y) = C(T)\, R(T, Y) \qquad \text{and} \qquad B(T, Y) = -C(T)\,\frac{Q'}{\gcd(Q, Q')} \tag{14}$$

with $C(T)$ ranging over $\mathbf{Q}(T) - \{0\}$.

Analogously we may consider the $\mathbf{Q}$-linear mapping φ_t from $\mathbf{Q}(t)^{r+1} \oplus \mathbf{Q}(t)^r$ to $\mathbf{Q}(t)^{d+r}$ given by the polynomial expression

$$a(Y)\, Q'(t, Y) + b(Y)\, Q(t, Y)$$

and the degree restrictions $\deg a \leqslant r$ and $\deg b \leqslant r-1$.

Again this $\mathbf{Q}$-linear mapping has a one-dimensional kernel which can be described by the polynomial expressions

$$a(Y) = cR(t, Y) \qquad \text{and} \qquad b(Y) = -c\,\frac{Q'(t, Y)}{\gcd(Q(t, Y), Q'(t, Y))} \tag{15}$$

with c ranging over $\mathbf{Q} - \{0\}$.

Since the linear mappings φ and φ_t have one-dimensional kernel and their source space is $2r+1$-dimensional, they have $2r$-dimensional image. With respect to the respective canonical vectorspace bases the matrices of φ and φ_t are suitable $(d+r)\times(2r+1)$ submatrices of the Sylvester matrices of $Q(T, Y)$, $Q'(T, Y)$ and $Q(t, Y)$, $Q'(t, Y)$ respectively. Since φ and φ_t have $2r$-dimensional images their matrices have both rank $2r$.

Selecting $2r$ linear independent rows from the matrix of φ_t we obtain a matrix $M_t \in \mathbf{Q}^{2r\times(2r+1)}$ of rank $2r$ which determines a homogeneous linear equation system admitting up to scaling just one nontrivial solution in $\mathbf{Q}^{2r+1}$.

From (15) we deduce that this nontrivial solution represents the coefficients of the univariate and monic polynomials $R(t, Y)$ and $Q'(t, Y)/\mathrm{gcd}(Q(t, Y), Q'(t, Y))$ up to a common nonzero rational scaling factor. The matrix M_t is a submatrix of the matrix of φ_t and hence of the Sylvester matrix of $Q(t, Y)$ and $Q'(t, Y)$. Therefore M_t is determined by a certain choice of $2r$ rows and $2r+1$ columns of the Sylvester matrix of $Q(t, Y)$ and $Q'(t, Y)$. By the same choice of rows and columns we determine a certain $2r\times(2r+1)$ submatrix $M(T)$ of the Sylvester matrix of $Q(T, Y)$ and $Q'(T, Y)$. Thus $M(T)\in\mathbf{Q}[T]^{2r\times(2r+1)}$.

Since the $2r\times(2r+1)$ matrix M_t has rank $2r$ we may choose a nonzero $2r\times 2r$-minor of it. The corresponding minor of $M(T)$ is nonzero too. Therefore we see that the rank of the matrix $M(T)\in\mathbf{Q}[T]^{2r\times(2r+1)}$ is $2r$ as well. In a similar way as before we deduce from (14) that the homogeneous linear equation system given by the matrix M determines the (in Y monic) polynomials $R(T, Y)$ and $Q'/\mathrm{gcd}(Q, Q')$ up to a nonzero scaling factor of $\mathbf{Q}(T)$. Moreover we have $M(t)=M_t$.

For $1\leqslant j\leqslant r+1$, let $M_j(T)$ denote the minor of $M(T)$ obtained by deleting the jth column of $M(T)$. Appending for each $1\leqslant i\leqslant n$ to the matrix $M(T)$ its ith row as the first row, and developing the determinant of this new matrix by its first row, we immediately deduce that the vector $(M_1(T), -M_2(T), ..., M_{2r+1}(T))$ is a nontrivial solution of the homogeneous linear equation system represented by the rank $2r$ matrix $M(T)$. Similarly we see that the vector $(M_1(t), -M_2(t), ..., M_{2r+1}(t))$ is also a nontrivial solution of the homogeneous linear equation system represented by the rank $2r$ matrix $M_t=M(t)$. Since the polynomial $R(t, Y)$ is monic we deduce easily from (15) that

$$M_1(t)\ Y^r - M_2(t)\ Y^{r-1} + \cdots + (-1)^r\, M_{r+1}(t) = M_1(t)\ R(t, Y)$$

and that therefore $M_1(t)\neq 0$ holds. This implies $M_1(T)\neq 0$. In an analogous way we deduce from (14) the identity

$$M_1(T)\ Y^r - M_2(T)\ Y^{r-1} + \cdots + (-1)^r\, M_{r+1}(T) = M_1(T)\ R(T, Y).$$

Taking into account $M_1(T) \neq 0$ we conclude

$$R(T, Y) := \frac{1}{M_1(T)} \sum_{0 \leqslant j \leqslant r} (-1)^{r-j} M_{r+1-j}(T)\, Y^j.$$

Now we are going to construct our approximation polynomial $\tilde{R}(T, Y)$.

Let $\tilde{M}(T)$ be the $2r \times (2r+1)$ submatrix of the Sylvester matrix of $\tilde{Q}$ and $\tilde{Q}' := \partial \tilde{Q}/\partial Y$ determined by the same choice of rows and columns as the matrices M_t and $M(T)$ (this makes sense because we have, by assumption, $\deg_Y \tilde{Q} = \deg_Y Q$).

Again, for $1 \leqslant j \leqslant r+1$, denote by $\tilde{M}_j(T)$ the minor of $M(T)$ obtained by deleting the jth column. As by assumption $\tilde{Q}(t, Y) = Q(t, Y)$ holds we deduce easily the identities $\tilde{M}_j(t) = M_j(t)$ for $1 \leqslant j \leqslant r+1$. In particular we have $\tilde{M}_1(t) \neq 0$ and $\tilde{M}_1(T) \neq 0$. More precisely, $M_1(t) \neq 0$ implies that $\tilde{M}_1(T)$ is a unit of the local $\mathbf{Q}$-algebra $\mathbf{Q}[T]_{\mathcal{M}}$. In particular, for any $1 \leqslant j \leqslant r+1$ the rational function $\tilde{M}_j(T)/\tilde{M}_1(T)$ belongs to $\mathbf{Q}[T]_{\mathcal{M}}$. We would like to use the expression

$$\frac{1}{\tilde{M}_1(T)} \sum_{0 \leqslant j \leqslant r} (-1)^{r-j} \tilde{M}_{r+1-j}(T)\, Y^j \tag{16}$$

for the definition of the approximation polynomial $\tilde{R}(T, Y)$.

However, the expression (16) contains terms which are rational functions in $T = (T_1, ..., T_m)$. In order to avoid this difficulty we replace the rational functions $\tilde{M}_{r+1-j}(T)/\tilde{M}_1(T)$ occurring in (16) by suitable approximation polynomials of $\mathbf{Q}[T]$. For this purpose we proceed as follows.

Without loss of generality we may assume that $M_1(t) = \tilde{M}_1(t) = 1$ holds. Let us write $H := 1 - M_1$ and $\tilde{H} := 1 - \tilde{M}_1$. Hence, $H(t) = \tilde{H}(t) = 0$ and we obtain the following identities in the power series ring $\mathbf{Q}[\![T - t]\!]$:

$$\frac{1}{M_1} = \sum_{\ell \geqslant 0} H^\ell \qquad \text{and} \qquad \frac{1}{\tilde{M}_1} = \sum_{\ell \geqslant 0} \tilde{H}^\ell.$$

Thus, in order to approximate $R(T, Y)$ with precision s in $\mathbf{Q}[\![T - t]\!]$, we define the polynomial $\tilde{R}(T, Y) \in \mathbf{Q}[T, Y]$ as

$$\tilde{R}(T, Y) := \sum_{0 \leqslant \ell \leqslant s} \tilde{H}^\ell(T) \sum_{0 \leqslant j \leqslant r} (-1)^{r-j} \tilde{M}_{r+1-j}(T)\, Y^j.$$

Let us verify that $\tilde{R}$ satisfies the requirements of the statement of Lemma 7:

- Obviously we have $\tilde{R} \neq 0$ and from $\tilde{M}_1 \neq 0$ we deduce that $\deg_Y \tilde{R} = r = \deg_Y R$ holds.

- Since by assumption the polynomial $\tilde{Q}$ approximates Q in $\mathbf{Q}[\![T-t]\!][Y]$ with precision s, we immediately see that for $1 \leqslant j \leqslant r+1$ the polynomial $\tilde{M}_j$ approximates M_j with precision s and that the polynomial $\sum_{0 \leqslant \ell \leqslant s} \tilde{H}^{\ell}$ approximates the power series $1/M_1 = \sum_{\ell \geqslant 0} H^{\ell}$ with precision s. This implies that the polynomial $\tilde{R}(T, Y)$ approximates $R(T, Y)$ with precision s in $\mathbf{Q}[\![T-t]\!][Y]$.

Finally, we design a suitable arithmetic network which produces the polynomial $\tilde{R}$ from the input circuit $\tilde{\beta}$ representing $\tilde{Q}$ and the coefficient representation of $Q(t, Y) = \tilde{Q}(t, Y)$.

The first step in the construction consists in writing down the matrix of the linear mapping φ_t and to extract $2r$ independent rows from it. This allows to identify the matrices $M_t = \tilde{M}(t)$ and $\tilde{M}(T)$ as submatrices of the Sylvester matrices of $Q(t, Y)$, $Q'(t, Y)$ and $\tilde{Q}(T, Y)$, $\partial \tilde{Q}/\partial Y(T, Y)$. This can be done by means of an arithmetic network of size $r^{O(1)}$ and nonscalar depth $O(\log r)$ using the coefficients of $Q(t, Y) = \tilde{Q}(t, Y)$ which represent part of the input. Then we compute the polynomials $\tilde{M}_j(T)$. This can be done by a division-free arithmetic circuit in $\mathbf{Q}[T]$ (which extends the input circuit $\tilde{\beta}$) of size $\tilde{L} + (2r+1)^{O(1)}$ and nonscalar depth $\tilde{\lambda} + O(\log(2r))$. Finally we compute $\tilde{H}$, $\sum_{0 \leqslant \ell \leqslant s} \tilde{H}^{\ell}$ and all the coefficients in Y of $\tilde{R}$ using $O(r \log s)$ additional arithmetic operations, organized in a circuit of nonscalar depth $O(\log s)$. In this way we obtain a division-free arithmetic circuit $\tilde{\gamma}$ of required size and nonscalar depth. ∎

Combining our previous results, we are now able to prove Theorem 2.

Proof of Theorem 2. Let $k := \lfloor \log_2 \deg_T P \rfloor + 1$.

Applying Lemma 5 to the input circuit β, we obtain a division-free arithmetic circuit γ_0 in $\mathbf{Q}[T, X]$ of size $O(kd^2n^7L)$ and non-scalar depth $O(k(\log_2 n + \lambda))$ which evaluates numerators $g_1 := g_1^{(k)}, \ldots, g_n := g_n^{(k)}$ and a non-zero denominator $h := h^{(k)}$ for the k-fold iteration of the Newton–Hensel operator N_F^k (such that $N_F^k(X) = (g_1/h, \ldots, g_n/h)$ holds).

Now, we apply γ_0 to compute the matrices

$$g_1(T, v_1(M), \ldots, v_n(M)), \ldots, g_n(T, v_1(M), \ldots, v_n(M))$$

and

$$h(T, v_1(M), \ldots, v_n(M))$$

in $\mathbf{Q}[T]^{D \times D}$ (compare with (4)). This yields an arithmetic circuit γ_1 in $\mathbf{Q}[T]$ of size $O(kd^2n^7D^2L)$.

By means of the well-parallelizable polynomial algorithm of Berkowitz [Ber84] we compute the coefficients of the characteristic polynomial $\mathcal{X}_h \in \mathbf{Q}[T][Y]$ of the matrix $h(T, v_1(M), ..., v_n(M))$.

Let

$$\mathcal{X}_h = \sum_{\ell=0}^{D} a_\ell(T)\, Y^\ell$$

with $a_0, ..., a_D \in \mathbf{Q}[T]$ and let

$$h^*(T, X) := \sum_{\ell=1}^{D} a_\ell(T)\, h^{\ell-1}.$$

Recall from the comments on Lemma 5 that the matrix $h(T, v_1(M), ..., v_n(M))$ is unimodular in $\mathbf{Q}[T]_{\mathcal{M}}^{D\times D}$. Thus a_0 is a unit of $\mathbf{Q}[T]_{\mathcal{M}}$ and in particular we have $a_0(t) \neq 0$.

Hence, the Cayley–Hamilton Theorem implies that

$$h(T, v_1(M), ..., v_n(M))^{-1} = \frac{-1}{a_0} h^*(T, v_1(M), ..., v_n(M)) \tag{17}$$

holds in $\mathbf{Q}[T]_{\mathcal{M}}^{D\times D}$. We compute now the entries of the matrix $h^*(T, v_1(M), ..., v_n(M))$ by means of an arithmetic circuit γ_2 which extends γ_1 increasing the size and nonscalar depth of γ_1 by $D^{O(1)}$ and $O(\log D)$, respectively.

We would like to compute the matrix $h(T, v_1(M), ..., v_n(M))^{-1}$ by means of formula (17). However, we are unable to do that in $\mathbf{Q}[T]$ because this would require division by the polynomial $a_0(T)$. Therefore we replace the matrix $h(T, v_1(M), ..., v_n(M))^{-1}$ by a suitable approximation in our subsequent argumentation. For this purpose let us use the following identity in $\mathbf{Q}[\![T-t]\!]$:

$$\frac{-1}{a_0(T)} = \frac{1}{a_0(t)} \sum_{\ell=0}^{\infty} \left(\frac{a_0(T) - a_0(t)}{a_0(t)}\right)^{\ell}.$$

Thus, the polynomial

$$a_0^*(T) := \frac{1}{a_0(t)} \sum_{\ell=0}^{2^k} \left(\frac{a_0(T) - a_0(t)}{a_0(t)}\right)^{\ell}$$

approximates the rational function $-1/a_0(T)$ in $\mathbf{Q}[\![T-t]\!]$ with precision 2^k.

We are now going to compute an approximation of the matrix

$$\tilde{M}_G := G\left(T, \frac{g_1}{h}(T, v_1(M), ..., v_n(M)), ..., \frac{g_n}{h}(T, v_1(M), ..., v_n(M))\right)$$

(recall that, by Lemma 6, the characteristic polynomial $\tilde{\mathscr{X}}_G$ of $\tilde{M}_G$ approximates the characteristic polynomal $\mathscr{X}_G$ of the homothety induced by G in $Q(T)[X]/(F_1, ..., F_n)$ with precision 2^k).

Let us consider the matrix

$$\hat{M}_G := G(T, (a_0^* h^* g_1)(T, v_1(M), ..., v_n(M)), ..., \\ (a_0^* h^* g_n)(T, v_1(M), ..., v_n(M))). \qquad (18)$$

The polynomial entries of the matrix $\hat{M}_G$ clearly approximate the corresponding entries of the matrix $\tilde{M}_G$ with precision 2^k in $\mathbf{Q}[\![T-t]\!]$. In order to compute $a_0^*(T)$ from $a_0(T)$, we use the division-free arithmetic circuit underlying the formula,

$$\sum_{\ell=0}^{2^k} \left(\frac{a_0(T)-a_0(t)}{a_0(t)}\right)^{\ell} \\ = \left(1+\frac{a_0(T)-a_0(t)}{a_0(t)}\right)\left(1+\left(\frac{a_0(T)-a_0(t)}{a_0(t)}\right)^2\right) \\ \cdots\left(1+\left(\frac{a_0(T)-a_0(t)}{a_0(t)}\right)^{2^{k-1}}\right) \\ +\left(\frac{a_0(T)-a_0(t)}{a_0(t)}\right)^{2^k}.$$

Plugging this circuit in the arithmetic circuit underlying formula (18) we obtain a division-free circuit γ_3 which extends γ_2. The arithmetic circuit γ_3 computes the entries of $\hat{M}_G$ and has the same asymptotic complexity as γ_2, namely size $O(kd^2n^7D^2L) + D^{O(1)}$ and nonscalar depth $O(k(\log n + \lambda) + \log D)$.

Observe that the characteristic polynomial $\hat{\mathscr{X}}_G$ of the matrix $\hat{M}_G$ approximates the characteristic polynomial $\tilde{\mathscr{X}}_G$ of $\tilde{M}_G(T, Y)$ with precision 2^k in $\mathbf{Q}[\![T-t]\!]$. Hence, $\hat{\mathscr{X}}_G$ also approximates the characteristic polynomial $\mathscr{X}_G$ of the homothety induced by G in $Q(T)[X]/(F_1, ..., F_n)$ with the same precision.

Applying the algorithm of Berkowitz [Ber84], we extend γ_3 to a circuit γ_4 which computes the coefficients of the characteristic polynomial $\hat{\mathcal{X}}_G$. The size and nonscalar depth of γ_4 are asymptotically of the same order than those of γ_3.

Now, observing that the polynomials $\hat{\mathcal{X}}_G$ and $\mathcal{X}_G$ satisfy the assumptions of Lemma 7, we obtain from the coefficients of $\hat{\mathcal{X}}_G$ of $\mathbf{Q}[T][Y]$ the coefficients of a polynomial $\tilde{P} \in \mathbf{Q}[T][Y]$ in Y which approximates with precision 2^k in $\mathbf{Q}[\![T-t]\!]$ the required minimal polynomial P of the homothety induced by G. The division-free arithmetic circuit γ_5 in $\mathbf{Q}[T]$ we use by means of Lemma 7 for this purpose, increases the size and depth of γ_4 by $kD^{O(1)}$ and $O(\log D + k)$ respectively. Thus, γ_5 has size $O(kd^2n^7D^2L) + kD^{O(1)}$ and nonscalar depth $O(k(\log n + \lambda) + \log D)$.

Finally, we compute the exact coefficients of P by the following variant of the Vermeidung von Divisionen technique.

In the following way we introduce a new variable Z in the arithmetic circuit γ_5 which evaluates the coefficients of the polynomial $\tilde{P} \in \mathbf{Q}[T][Y]$ with respect to the variable Y: we replace the variables $T_1, ..., T_m$ in γ_5 by the monomials $ZT_1, ..., ZT_m$. In this manner, we compute the coefficients of the polynomial

$$\tilde{P}(ZT_1, ..., ZT_m, Y) \in \mathbf{Q}[T, Z][Y]$$

with respect to the variable Y by means of a new division-free circuit γ_6 in $\mathbf{Q}[T, Z]$. The size and nonscalar depth of the circuit γ_6 are asymptotically the same as those of the circuit γ_5.

Now we modify the circuit γ_6 as follows: in each computation step we write the corresponding intermediate result as a polynomial in the variable Z (performing interpolation in Z) and eliminate all monomial terms of degree strictly greater than $\deg_T P$. Hence, simulating the execution of the arithmetic circuit γ_6 performing step by step the above modification, the elements of $\mathbf{Q}[T]$ we obtain as coefficients of monomials in Z are all the coefficients of $\tilde{P}$ up to degree $\deg_T P$.

In this way, we obtain a division-free arithmetic circuit γ_7 in $\mathbf{Q}[T, Z]$ which computes approximations to the input polynomials of γ_6 with precision $\deg_T P$ with respect to the variable Z. The final circuit γ is obtained from γ_7 by specializing the variable Z into the value 1.

Thus, γ computes the coefficients of the minimal polynomial P *exactly*. Finally let us notice that γ has asymptotically the same size and nonscalar depth than the circuit γ_7, and γ_7 has the same nonscalar depth than γ_6 whereas its size becomes increased by a factor $\deg_T^2 P$ with respect to the size of the circuit γ_6. Thus, the circuit γ has size $O(d^2n^7D^2 \deg_T^2 P \log(\deg_T P)\, L) + D^{O(1)} \deg_T^2 P \log(\deg_T P)$ and nonscalar depth $O((\log n + \lambda) \log(\deg_T P) + \log D)$. ▮

3. REFINEMENT OF THE MAIN RESULT AND EXAMPLES

3.1. *Refinement of the Main Result*

The complexity bound stated in our main result, namely Theorem 2, depends substantially on linear algebra computations with $D \times D$ matrices. The original $D \times D$ matrix we deal with is M, the companion matrix of the polynomial q appearing in the given description of a geometric solution of the fiber $\pi^{-1}(t)$. Suppose now that the Galois group of $\mathbf{C}$ over $\mathbf{Q}$ does *not* act transitively on the fiber $\pi^{-1}(t)$ and that a certain factorization of the polynomial q over $\mathbf{Q}[Y]$ is given, capturing this situation. Let us suppose that q has the form

$$q = \prod_{j=1}^{s} q_j,$$

where $q_1, ..., q_s$ are nonconstant polynomials of $\mathbf{Q}[Y]$ with $D_j := \deg q_j$ and $\sum_{j=1}^{s} D_j = D$. Observe that the polynomials $q_1, ..., q_s$ are pairwise coprime since the polynomial q is separable by assumption.

Instead of working with the companion matrix M directly, we may take into account its special block form corresponding to the given factorization of the polynomial q. Thus, let us consider the matrix

$$M^* := \begin{pmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_s \end{pmatrix},$$

where, for each $1 \leqslant j \leqslant s$, we denote by M_j the companion matrix of the polynomial q_j.

From the fact that the polynomials $q_1, ..., q_s$ are pairwise coprime and the Chinese remainder theorem, we deduce easily that M^* is similar to the matrix M. Thus q is the characteristic and—by its separability—also the minimal polynomial of M^*.

As it is unlikely that linear algebra routines dealing with special matrices like those occurring in polynomial elimination theory will become one day linear time procedures, we may profit from the block structure of the matrix M^*.

For example, if during our main algorithm (see proof of Theorem 2), we need to apply a quadratic time linear algebra routine to the $D \times D$ matrix M, replacing the matrix M by the similar matrix M^* the complexity becomes $O(D_1^2 + \cdots + D_s^2)$ instead of $O(D^2) = O((D_1 + \cdots + D_s)^2)$. Doing so, the time economy is of order $O(\sum_{1 \leqslant j < k \leqslant s} D_j D_k)$, which may be of decisive importance when implementing (and running) our main algorithm on a real world computer.

This leads to the following reformulation of our main result in the case that a factorization of the polynomial q in nonconstant polynomials $q_1, ..., q_s \in \mathbf{Q}[Y]$ is given.

COROLLARY 8. *Let notations and assumptions be as in Theorem 2. Assume, furthermore, that there are given s coprime polynomials $q_1, ..., q_s \in \mathbf{Q}[Y]$ of positive degrees $D_1, ..., D_s$ respectively such that $q = q_1 \cdot \cdots \cdot q_s$ holds. Let ω be the maximal exponent of the matrix operations we are going to use during our algorithm. Then, there exists a division-free arithmetic circuit γ in $\mathbf{Q}[T]$ with the following properties*

(i) *γ computes the coefficients of the polynomial P with respect to the main variable Y*

(ii) *γ has size $O(d^2 n^7 (D_1^2 + \cdots + D_s^2) \deg_T^2 P \log(\deg_T P)\, L + (D_1^\omega + \cdots + D_s^\omega) \deg_T^2 P \log(\deg_T P))$*

(iii) *γ has nonscalar depth $O((\log n + \lambda) \log(\deg_T P) + \log(\max\{D_i;\ 1 \leqslant i \leqslant s\}))$*

Moreover, there exists a uniform family of arithmetic networks of asymptotically the same size and nonscalar depth than γ which produces γ from the following data:

— *the input arithmetic circuit β*

— *the rational numbers which represent the coordinates of t and the given description of a geometric solution of the fiber $\pi^{-1}(t)$ (containing the polynomial q)*

— *the coefficients of the polynomials $q_1, ..., q_s$.*

3.2. *Examples*

EXAMPLE 1. The aim of this example is to analyze what occurs with the following parametric situation when our main result is applied.

Let $T := (T_1, ..., T_m)$ be indeterminates and let $A_{n-1}, ..., A_0$ be polynomials of $\mathbf{Q}[T]$ of degree not exceeding an a priori fixed bound Δ. Suppose that the polynomials $A_{n-1}, ..., A_0$ are given by a division-free arithmetic circuit β in $\mathbf{Q}[T]$. Let L be the size of β.

We consider the following polynomials of $\mathbf{Q}[T][X]$

$$F_{n-1} := \sigma_{n-1} + A_{n-1}, ..., F_0 := \sigma_0 - (-1)^n A_0$$

(where σ_i denotes the ith elementary symmetric function in the variables $X := (X_1, ..., X_n)$). Obviously, the zeroes of the polynomial equation system

$$F_0 = 0, ..., F_{n-1} = 0 \tag{19}$$

are closely related to the zeroes of the polynomial equation $Q = 0$ with

$$Q(Y) = Y^n + A_{n-1} Y^{n-1} + \cdots + A_0 \in \mathbf{Q}[T][Y].$$

First, let us analyze the non-parametric case of the equation system (19).

For given $a_0, ..., a_{n-1} \in \mathbf{Q}$, let us consider the following polynomials of $\mathbf{Q}[X]$:

$$f_{n-1} := \sigma_{n-1} + a_{n-1}, ..., f_0 := \sigma_0 - (-1)^n a_0.$$

Assume that the corresponding polynomial $Q^* := Y^n + a_{n-1} Y^{n-1} + \cdots + a_0 \in \mathbf{Q}[Y]$ is *separable*.

The polynomials $f_{n-1}, ..., f_0$ define a zero-dimensional equation system without parameters (corresponding to the polynomial equation system (19) with $m = 0$) which admits $n!$ solutions. Observe that this system has no solution at infinity and that its Bézout-number is also $n!$.

Observe that the polynomials $Q^*(X_1), ..., Q^*(X_n)$ represent a succinct uncoupling of the variables $X_1, ..., X_n$ occurring in the polynomial system $f_{n-1} = 0, ..., f_0 = 0$. The polynomials $Q^*(X_1), ..., Q^*(X_n)$ can be represented by an array of n^2 elements of $\mathbf{Q}$ or by a division-free arithmetic circuit of length $O(n \log n)$ in $\mathbf{Q}[X]$.

However, for the input polynomials $f_0, ..., f_{n-1}$, the main algorithm of [GHH$^+$97] and [GHMP97] (like many other symbolic procedures designed for the special purpose of algorithmic Galois theory as, e.g., [Val95, Col97]) chooses a sufficiently generic $\mathbf{Q}$-linear combination U of the variables $X_1, ..., X_n$ and produces the resolvent, say R_U, of the equation $Q = 0$ with respect to the variable U. The resolvent R_U belongs to the polynomial ring $\mathbf{Q}[U]$ and its degree equals the order of the Galois group of the splitting field of Q. The main algorithm of [GHH$^+$97, GHM$^+$98] requires $(n!)^{O(1)} = n^{O(n)}$ arithmetic operations in $\mathbf{Q}$ in order to produce the polynomial R_U.

Let us now come back to the situation of the original parametric equation system $F_0 = 0, ..., F_{n-1} = 0$.

Let us suppose that the polynomial Q is separable with respect to the variable Y. Note that the polynomials $F_{n-1}, ..., F_0$ form a regular sequence in $\mathbf{Q}[T, X]$ and, therefore, they define an equidimensional variety of dimension m, namely

$$V := \{F_{n-1} = 0, ..., F_0 = 0\} \subset \mathbf{A}^{m+n}.$$

The morphism of affine varieties $\pi\colon V \rightarrow \mathbf{A}^m$ induced by the canonical projection of $\mathbf{A}^{m+n}$ onto $\mathbf{A}^m$ is finite, generically unramified and has degree $n!$.

Suppose that there is given a point $t \in \mathbf{Q}^m$ for which the specialized polynomial $Q(t, Y) := Y^n + A_{n-1}(t)\, Y^{n-1} + \cdots + A_0(t)$ is separable (i.e., where the fiber $\pi^{-1}(t)$ is unramified).

Assume furthermore that we have "sufficient information" about the geometric nature of the fiber $\pi^{-1}(t)$. In particular, we suppose that the polynomial $Q(t, Y)$ (and the Galois group of $Q(t, Y)$) is known. Then, applying the arithmetic network of Theorem 2 to the general problem instance $F_0, ..., F_{n-1}, G$ with $G := X_1$, we are able to find the polynomial Q in sequential time $O(n^9 L\Delta^2 \log \Delta + (n!)^{\omega}\, \Delta^2 \log \Delta)$ where ω denotes the maximal exponent of matrix operations used in the algorithm. In case $Q(t, Y)$ has a known decomposition into linear factors of $\mathbf{Q}[Y]$, Corollary 8 implies this complexity bound can be lowered to $O(n^9 L\Delta^2 \log \Delta + n!\, \Delta^2 \log \Delta)$. But in fact, in this particular case, it is possible to design a variant of our algorithm with better complexity bounds:

Remark 9. Let notations be as before and assume that the univariate polynomial $Q(t, Y)$ has a known decomposition into linear factors of $\mathbf{Q}[Y]$. Then there exists a division-free arithmetic circuit γ in $\mathbf{Q}[T]$ of size $O(n^9 L\Delta^2 \log \Delta)$ and nonscalar depth $O(\log \Delta(\log n + \lambda))$ which recovers the polynomial $Q \in \mathbf{Q}[T, Y]$.

Proof. Let $y_1, ..., y_n \in \mathbf{Q}$ be the n distinct roots of $Q(t, Y)$. As we want to compute the minimal polynomial of $G := X_1$ modulo the ideal $(F_0, ..., F_{n-1})$ and since $\pi^{-1}(t) = \{t\} \times V_t$ with

$$V_t = \{\tau(y_1, ..., y_n) : \tau \text{ is a permutation of } n \text{ elements}\}$$

holds, it suffices to consider the n distinct rational points

$$\begin{aligned} \xi^{(1)} &:= (y_1, y_2, ..., y_n) \\ \xi^{(2)} &:= (y_2, ..., y_n, y_1) \\ &\vdots \\ \xi^{(n)} &:= (y_n, y_1, ..., y_{n-1}) \end{aligned}$$

all belonging to V_t and having distinct first coordinates.

Applying Lemma 3 to the rational points $\xi^{(1)}, ..., \xi^{(n)} \in V_t$ we obtain for each $1 \leqslant \ell \leqslant n$ an n-tuple of power series $R^{\xi^{(\ell)}} := (R_1^{(\xi^{(\ell)})}, ..., R_n^{(\xi^{(\ell)})}) \in \mathbf{Q}[\![T-t]\!]^n$ such that $R^{\xi^{(\ell)}}(t) = \xi^{\ell}$ and $F_0(T, R^{\xi^{(\ell)}}) = 0, ..., F_{n-1}(T, R^{\xi^{(\ell)}}) = 0$ holds.

Since $R_1^{(\xi^{(\ell)})}(t) = y_\ell \neq y_{\ell'} = R_1^{\xi^{(\ell')}}$ holds for any $1 \leqslant \ell \neq \ell' \leqslant n$, we conclude that the series $R_1^{\xi^{(1)}}, ..., R_1^{\xi^{(n)}} \in \mathbf{Q}[\![T-t]\!]$ are all distinct. Hence the minimal polynomial of the homothety η_{X_1} of multiplication by X_1 in $Q(T)[X]/(F_0, ..., F_{n-1})$ can be written as $P_{X_1} := \prod_{\ell=1}^{n} (Y - R_1^{(\xi^{(\ell)})})$.

Now, we repeat, one by one, exactly the same steps than in the proof of Theorem 2 in order to compute the minimal polynomial P_{X_1} which is the characteristic polynomial of the $n \times n$ matrix

$$\begin{pmatrix} R_1^{(\xi^{(1)})} & & \\ & \ddots & \\ & & R_1^{(\xi^{(n)})} \end{pmatrix}.$$

The reduction of the complexity of the algorithm underlying the statement of Remark 9 is simply due to the fact that instead of dealing with $n! \times n!$ matrices as in the proof of Theorem 2, we may restrict ourselves to the consideration of matrices having size only $n \times n$. ∎

Example 2. Let $T := (T_1, ..., T_n)$ and $X := (X_1, ..., X_n)$.

Let be given polynomials $f_1, ..., f_n \in \mathbf{Q}[X]$ of degree at most d which define a morphism $\varphi := (f_1, ..., f_n)$ of the affine space $\mathbf{A}^n$ onto itself which is supposed to be an automorphism of $\mathbf{A}^n$.

Suppose that $f_1, ..., f_n$ are represented by a division-free arithmetic circuit in $\mathbf{Q}[X]$ of length L.

Let $Q_1, ..., Q_n$ be the polynomials of $\mathbf{Q}[X]$ which define the inverse map of φ, i.e., let $\varphi^{-1} = (Q_1, ..., Q_n)$. Furthermore, let $\Delta := \max\{\deg Q_1, ..., \deg Q_n\}$.

We consider the polynomials $F_1 := T_1 - f_1, ..., F_n := T_n - f_n \in \mathbf{Q}[T, X]$.

The polynomials $F_1, ..., F_n$ generate a complete intersection ideal in $\mathbf{Q}[T, X]$ which is prime and hence radical. Hence, the variety V defined by the equations $F_1 = 0, ..., F_n = 0$ is equidimensional, of dimension n, and the morphism π: $V \to \mathbf{A}^n$ induced by the canonical first projection of $\mathbf{A}^{2n}$ onto $\mathbf{A}^n$, is an isomorphism of varieties. Therefore the morphism π is finite and unramified. Note that Bézout's Theorem implies that $\deg V \leqslant d^n$ holds. Moreover, we have $\deg \pi = 1$. From Gabber's Theorem [BCW82] we deduce $\Delta \leqslant d^n$.

Let $P_1, ..., P_n$ be the elimination polynomials of the linear forms $X_1, ..., X_n$ with respect to the variety V. These polynomials belong to $\mathbf{Q}[T][Y]$ and have the form $P_1 = Y - Q_1, ..., P_n = Y - Q_n$. Thus, the computation of

$P_1, \ldots, P_n$ yields the inverse map φ^{-1}. Let $F := (F_1, \ldots, F_n)$ and observe that the Jacobian matrix

$$DF(T, X) := \begin{pmatrix} \dfrac{\partial F_1(T, X)}{\partial X_1} & \cdots & \dfrac{\partial F_1(T, X)}{\partial X_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_n(T, X)}{\partial X_1} & \cdots & \dfrac{\partial F_n(T, X)}{\partial X_n} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1(X)}{\partial X_1} & \cdots & \dfrac{\partial f_1(X)}{\partial X_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n(X)}{\partial X_1} & \cdots & \dfrac{\partial f_n(X)}{\partial X_n} \end{pmatrix}$$

is unimodular, since φ is an automorphism of $\mathbf{A}^n$. This implies that $\det DF(T, X)$ is a nonzero rational number.

Using now the algorithm underlying Theorem 2 in order to compute the polynomials $P_1, \ldots, P_n$ and the fact that $\det DF(T, X)$ is a nonzero rational number, we obtain a division-free arithmetic circuit γ in $\mathbf{Q}[T]$ of size $O(Ln^7 d^2 \Delta^2 \log \Delta) = O(Ln^8 d^{2n+2} \log d) = O(Ln^8 d^{2n+3})$, which represents the polynomials $Q_1, \ldots, Q_n$ and, hence, the inverse map φ^{-1}.

REFERENCES

[AS88] W. Auzinger and H. J. Stetter, An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations, *in* "International Series of Numerical Mathematics," Vol. 80, pp. 11–30, Birkhäuser, Basel, 1988.

[BCA97] P. Bürgisser, M. Clausen, and M. Amin Shokrollahi, "Algebraic Complexity Theory," Grundlehren der Mathematischen Wissenschaften, Vol. 315, Springer-Verlag, New York/Berlin, 1997.

[BCW82] H. Bass, E. Connelle, and T. Wright, The Jacobian conjecture: Reduction of degree and formal expansion of the inverse, *Bull. Amer. Math. Soc.* **7** (1982), 287–330.

[Ber84] S. J. Berkowitz, On computing the determinant in small parallel time using a small number of processors, *Inform. Process. Lett.* **18** (1984), 147–150.

[BGHM] B. Bank, M. Giusti, J. Heintz, and G. Mbakop, Polar varieties and efficient real equation solving: The hypersurface case, *J. Complexity* **13** (1997), 5–27.

[Bor93] A. Borodin, Time space tradeoffs (getting closer to the barrier?), *in* "Algorithms and Computation, Proceedings 4, ISSAC," Lecture Notes in Comput. Sci., Vol. 762, pp. 209–220, Springer-Verlag, New York/Berlin, 1993.

[BPR97] S. Basu, R. Pollack, and M.-F. Roy, On computing a set of points meeting every cell defined by a family of polynomials on a variety, *J. Complexity* **13** (1997), 28–37.

[Bru98] N. Bruno, "Esquemas de compilación de circuitos aritméticos uniformes descriptos por medio de funciones generatrices," Master's thesis, FaMAF, Universidad de Córdoba, Argentina, 1998.

[Buc85] B. Buchberger, Gröbner bases: An algorithmic method in polynomial ideal theory, *in* "Multidimensional System Theory" (N. K. Bose *et al.*, Eds.), pp. 374–383, Reidel, Dordrecht, 1985.

[Can90] J. Canny, Generalised characteristic polynomials, *J. Symbolic Comput.* **9** (1990), 241–250.

[Cas97] D. Castro, "Sobre la complejidad de la aproximación diofántica y los fundamentos del análisis numérico," Master's thesis, Universidad de Cantabria, Santander, Spain, 1997.

[CG83] A. L. Chistov and D. Y. Grigoriev, Subexponential time solving systems of algebraic equations, LOMI preprint E-9-83, E-10-83, Steklov Institute, Leningrad, 1983.

[Col97] A. Colin, "Théorie des invariants effective," Ph.D. thesis, Ecole Polytechnique, Palaiseau-Paris, France, 1997.

[Ful84] W. Fulton, "Intersection Theory," 2nd ed., Ergebnisse der Mathematik, Springer-Verlag, New York/Berlin, 1984.

[GH93] M. Giusti and J. Heintz, La détermination des points isolés et de la dimension d'une variété algébrique peut se faire en temps polynomial, *in* "Computational Algebraic Geometry and Commutative Algebra, Symposia Mathematica" (D. Eisenbud and L. Robbiano, Eds.), Vol. 34, pp. 216–256, Cambridge Univ. Press, Cambridge, UK, 1993.

[GHH+97] M. Giusti, K. Hägele, J. Heintz, J. E. Morais, J. L. Montaña, and L. M. Pardo, Lower bounds for diophantine approximation, *J. Pure Appl. Algebra* **117–118** (1997), 277–317.

[GHL+97] M. Giusti, K. Hägele, G. Lecerf, J. Marchand, and B. Salvy, "Computing the Dimension of a Projective Variety: The Projective Noether Maple Package," Research Report 3224, INRIA, France, July 1997.

[GHM+98] M. Giusti, J. Heintz, J. E. Morais, J. Morgenstern, and L. M. Pardo, Straight-line programs in geometric elimination theory, *J. Pure Appl. Algebra* **124** (1998), 101–146.

[GHMP97] M. Giusti, J. Heintz, J. E. Morais, and L. M. Pardo, Le rôle des structures de données dans les problèmes d'élimination, *C. R. Acad. Sci. Paris* **325** (1997), 1223–1228.

[GM89] P. Gianni and T. Mora, Algebraic solution of systems of polynomial equations using Gröbner bases, *in* "Proceedings AAECC-5," Lecture Notes in Comput. Sci., Vol. 356, pp. 247–257, Springer-Verlag, New York/Berlin, 1989.

[GV88] D. Yu. Grigoriev and N. N. Vorobjov, Jr., Solving systems of polynomial inequalities in sub-exponential time, *J. Symbolic Comput.* **5** (1988), 37–64.

[Hei83] J. Heintz, Definability and fast quantifier elimination in algebraically closed fields, *Theoret. Comput. Sci.* **24** (1983), 239–277.

[HMPS98] K. Hägele, J. E. Morais, L. M. Pardo, and M. Sombra, On the intrinsic complexity of the arithmetic Nullstellensatz, *J. Pure Appl. Algebra*, in press.

[HRS90] J. Heintz, M.-F. Roy, and P. Solernó, Sur la complexité du principe de Tarski–Seidenberg, *Bull. Soc. Math. France* **118** (1990), 101–126.

[Ive73] B. Iversen, Generic local structure of the morphisms in commutative algebra, *in* "Lecture Notes in Math.," Vol. 310, Springer-Verlag, New York/Berlin, 1973.

[Kal85] E. Kaltofen, Computing with polynomials given by straight-line programs. I. Greatest common divisors, *in* "Proceedings of the 17th Ann. ACM Symposium on Theory of Computing, Providence, RI," pp. 131–142, ACM Press, New York, 1985.

[Koe03] J. Koenig, "Einleitung in die allgemeine Theorie der algebraischen Grössen," Teubner, Leipzig, 1903.

[Kov90] P. Kovacs, Minimum degree solutions for the inverse kinematics problem by application of the Buchberger algorithm, *in* "Proceedings of the 2nd Int. Workshop on Advances in Robot Kinematics, University of Linz, Austria, September 10–12, 1990," Springer-Verlag, New York/Berlin, 1990.

[KP96] T. Krick and L. M. Pardo, A computational method for diophantine approximation, *in* "Algorithms in Algebraic Geometry and Applications, Proceedings of MEGA'94" (L. González-Vega and T. Recio, Eds.), Progress in Mathematics, Vol. 143, pp. 193–254, Birkhäuser Verlag, Basel, 1996.

[Kro82] L. Kronecker, Grundzüge einer arithmetischen Theorie de algebraischen Grössen, *J. Reine Angew. Math.* **92** (1882), 1–22.

[LV98] M. J. Gonzalez Lopez and L. Gonzalez Vega, Newton identities in the multivariate case: Pham systems, *in* "London Mathematical Society Lecture Notes Series," Vol. 251, pp. 351–366, Cambridge Univ. Press, Cambridge, UK, 1998.

[Mac16] F. S. Macaulay, "Algebraic Theory of Module Systems," Cambridge Univ. Press, Cambridge, UK, 1916.

[Mat97] G. Matera, "Sobre la complejidad en espacio y tiempo de la eliminación geométrica," Ph.D. thesis, Universidad de Buenos Aires, Argentina, 1997.

[Mis93] B. Mishra, "Algorithmic Algebra," Springer-Verlag, New York, 1993.

[Mor97] J. E. Morais, "Resolución eficaz de sistemas de ecuaciones polinomiales," Ph.D. thesis, Universidad de Cantabria, Santander, Spain, 1997.

[MP97] B. Mourrain and V. Pan, Solving special polynomial systems by using structural matrices and algebraic residues, *in* "Proceedings Foundations of Computational Mathematics, FOCM'97" (F. Cucker and M. Shub, Eds.), pp. 287–304, Springer-Verlag, New York/Berlin, 1997.

[Nar66] R. Narasimhan, Introduction to the theory of analytic spaces, *in* "Lecture Notes in Math.," Vol. 25, Springer-Verlag, New York/Berlin, 1966.

[Ren92] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. Part I. Introduction, preliminaries, the geometry of semi-algebraic sets, the decision problem for the existential theory of the reals, *J. Symbolic Comput.* **13** (1992), 255–299.

[Sch95] J. Schmid, On the affine Bézout inequality, *Manuscripta Math.* **88** (1995), 225–232.

[SGV94] A. Schönhage, F. W. Grotefeld, and E. Vetter, "Fast Algorithms: A Multitape Turing Machine Implementation," Wissenschaftsverlag, Mannheim, 1994.

[Som97] M. Sombra, Bounds for the Hilbert function of polynomial ideals and for the degrees in the Nullstellensatz, *J. Pure Appl. Algebra* **117–118** (1997), 565–599.

[SS93a] M. Shub and S. Smale, Complexity of Bézout's theorem. I. Geometric aspects, *J. Amer. Math. Soc.* **6** (1993), 459–501.

[SS93b] M. Shub and S. Smale, Complexity of Bézout's theorem. II. Volumes and probabilities, *in* "Proceedings of MEGA'92," Progress in Mathematics, Vol. 109, pp. 267–285, Birkhäuser Verlag, Basel, 1993.

[SS93c] M. Shub and S. Smale, Complexity of Bézout's theorem. III. Condition number and packing, *J. Complexity* **9** (1993), 4–14.

[SS94] M. Shub and S. Smale, Complexity of Bézout's theorem. V. Polynomial time, *Theoret. Comput. Sci.* **133** (1994), 141–164.

[SS96a] J. Sabia and P. Solernó, Bounds for traces in complete intersections and degrees in the Nullstellensatz, *Appl. Algebra Engrg. Comm. Comput.* **6** (1996), 353–376.

[SS96b] M. Shub and S. Smale, Complexity of Bezout's theorem. IV. Probability of success and extensions, *SIAM J. Numer. Anal.* **33** (1996), 128–148.

[Str73] V. Strassen, Vermeidung von Divisionen, *J. Reine Angew. Math.* **264** (1973), 182–202.

[TER97] TERA Development Group, Some remarks on the time-space tradeoff of geometric elimination procedures, manuscript, 1997.

[Val95] A. Vallibouze, Computations of the Galois groups of the resolvent factors for the direct and indirect Galois problem, *in* "Proceedings of AAECC-11" (M. Giusti, G. Cohen, and T. Mora, Eds.), Lecture Notes in Comput. Sci., Vol. 948, pp. 456–468, Springer-Verlag, New York/Berlin, 1995.

[VH94] J. Verschelde and A. Haegemans, Homotopies for solving polynomial systems within a bounded domain, *Theoret. Comput. Sci.* **133** (1994), 165–185.

[Vog84] W. Vogel, "Results on Bezout's Theorem," Tata Institute of Fundamental Research, Springer-Verlag, New York/Berlin, 1984.

[vzG86] J. von zur Gathen, Parallel arithmetic computations: A survey, *in* "Proceedings of the 12th Symposium on Mathematical Foundations of Computer Science" (B. Rovan, J. Gruska, and J. Wiedermann, Eds.), Lecture Notes in Comput. Sci., Vol. 233, pp. 93–112, Springer-Verlag, New York/Berlin, 1986.

[vzG93] J. von zur Gathen, Parallel linear algebra, *in* "Synthesis of Parallel Algorithms" (J. H. Reif, Ed.), pp. 573–617, San Mateo, CA, 1993.

[WB93] V. Weispfenning and T. Becker, "Groebner Bases: A Computational Approach to Commutative Algebra," Graduate Texts in Mathematics, Vol. 141, Springer-Verlag, New York/Berlin, 1993.

[Zar95] O. Zariski, "Algebraic Surfaces," Classics in Mathematics, Springer-Verlag, New York/Berlin, 1995.