

THE CLASSICAL ARTIN APPROXIMATION THEOREMS

HERWIG HAUSER

To Michael Artin

ABSTRACT. The various Artin approximation theorems assert the existence of power series solutions of a certain quality Q (i.e., formal, analytic, algebraic) of systems of equations of the same quality Q , assuming the existence of power series solutions of a weaker quality $Q' < Q$ (i.e., approximated, formal). The results are frequently used in commutative algebra and algebraic geometry. We present a systematic argument which proves, with minor modifications, all theorems simultaneously. More involved results, such as, e.g., Popescu’s nested approximation theorem for algebraic equations or statements about the Artin function, will only be mentioned but not proven. We complement the article with a brief account of the theory of algebraic power series, two applications of approximation to singularities, and a differential-geometric interpretation of Artin’s proof.

CONTENTS

1. What is the problem?	595
2. The various approximation theorems	597
3. Proof of the analytic and algebraic approximation theorems	603
4. Proofs of the parametrization theorem and the four strong approximation theorems	607
5. Weierstrass division theorems	612
6. Algebraic power series	615
7. Formal and analytic relations between convergent series	619
8. Two applications of approximation	621
9. The geometry behind Artin’s proof	624
Acknowledgments	629
About the author	629
References	629

1. WHAT IS THE PROBLEM?

Let $f(t, y_1, \dots, y_m)$ be a polynomial in one distinguished variable t and m other variables y_1, \dots, y_m , with coefficients in a field \mathbb{K} . Assume that we wish to find all power series solutions $y(t) = (y_1(t), \dots, y_m(t))$ of the equation $f(t, y) = 0$. Expand $y_j(t) = \sum_{i=0}^{\infty} y_{ij} t^i$ as a series in t with unknown coefficients y_{ij} , for $j = 1, \dots, m$

Received by the editors December 2, 2016.

2010 *Mathematics Subject Classification*. Primary 13-02, 14-02, 32-02.

Supported by the Austrian Science Fund FWF within the projects P25652 and AI0038211.

and $i \geq 0$. Substitution of $y(t)$ into $f(t, y)$ yields a series $f(t, y(t))$ in t . Setting in this series the coefficients of t^ℓ equal to 0 we get, for $\ell = 0, 1, 2, \dots$, the polynomial equations

$$\begin{aligned} F_0(y_{01}, \dots, y_{0m}) &= 0, \\ F_1(y_{01}, \dots, y_{0m}, y_{11}, \dots, y_{1m}) &= 0, \\ &\dots \\ F_\ell(y_{01}, \dots, y_{0m}, \dots, y_{\ell 1}, \dots, y_{\ell m}) &= 0, \end{aligned}$$

each depending on the first ℓ m -tuples of the coefficients y_{ij} . Of course, F_0 coincides with f . The successive other polynomials F_ℓ are obtained from the preceding $F_{\ell-1}$ by applying to $F_{\ell-1}$ the universal derivation Δ of the polynomial ring $\mathbb{K}[y_{ij}]$ in countably many variables y_{ij} . It is given by applying the shifts $\Delta(y_{ij}) := y_{i+1,j}$ to the variables.

The natural approach to solving the above infinite system of equations in the y_{ij} is to proceed stepwise, first solving F_0 for y_{01}, \dots, y_{0m} , getting a set of solutions $\alpha_0 = (\alpha_{01}, \dots, \alpha_{0m}) \in \mathbb{K}^m$. Any choice of α_0 can be plugged into F_2 for y_{01}, \dots, y_{0m} , and results in the polynomial equation

$$F_1(\alpha_{01}, \dots, \alpha_{0m}, y_{11}, \dots, y_{1m}) = 0$$

for y_{11}, \dots, y_{1m} . The solvability of this second equation may depend on the prior choice of α_0 . Assume we have taken a good choice, so that we are able to find a solution $\alpha_1 = (\alpha_{11}, \dots, \alpha_{1m}) \in \mathbb{K}^m$ to the last equation. Now the game repeats. We wish to solve the next equation $F_2 = 0$ for α_2 . But the solvability of $F_2(\alpha_0, \alpha_1, y_{21}, \dots, y_{2m}) = 0$ may depend not only on the choice of α_1 , but possibly also on the earlier choice of α_0 . So things become more complicated.

One of Artin’s theorems (cited as “strong approximation theorem” in the literature) implies that for every f there is a bound, say e , so that if one can solve F_0, F_1, \dots up to F_e , then all remaining equations F_ℓ can be solved automatically. One says in this case that an approximate solution can be lifted to an exact solution. This is quite surprising, since it asserts that from a certain index on there is no more obstruction to solving F_ℓ . A more precise version says that, for any given c , there exists a bound e so that any solution $\alpha = (\alpha_{ij}, i \leq e, j = 1, \dots, m)$ for F_0, \dots, F_e has to be modified only in the components α_{ij} with $i \geq c$ in order to then admit a lifting to a solution of all F_ℓ . In general, a given solution α for F_0, \dots, F_e cannot be lifted directly: one has to go back to a smaller index c , take $\alpha' = (\alpha_{ij}, i \leq c, j = 1, \dots, m)$ and lift it to a solution for all F_ℓ . So the values of α_{ij} may have to be changed for the indices i between $c + 1$ and e in order to construct the solution $y(t)$ of $f(t, y) = 0$.

Of course, the formal solution $y(t)$ of $f(t, y) = 0$ is then given by the components $y_i(t) = \sum_{i=0}^\infty \alpha_{ij} t^i \in \mathbb{K}[[t]]$ for the properly chosen coefficients α_{ij} .

It is a priori not clear whether a solution constructed in this way is convergent, or whether one can modify a divergent solution so as to get a convergent solution. One has to understand how divergence arises. Does it arise by “mistake”? Can one always achieve convergence by smart choices of the values of α_{ij} ?

To answer this, observe that the solutions $\alpha_{i1}, \dots, \alpha_{im}$ chosen in the i th step are not unique. Indeed, one may choose excessively large values for the α_{ij} so that one still gets a solution of

$$F_i(\alpha_0, \alpha_1, \dots, \alpha_{i-1}, y_{i1}, \dots, y_{im}) = 0.$$

Said differently, one picks a point of the algebraic subvariety of \mathbb{K}^m defined by this equation which is very far away from the origin. Doing this again and again, one may find a formal solution $y(t)$ which is not convergent.

Artin’s analytic approximation theorem now asserts that these mistakes can always be avoided: Whenever the system $F_\ell(y) = 0$ admits a solution α (possibly giving rise to a formal solution $\widehat{y}(t)$ of $f(t, y) = 0$), one may also find a solution $\widetilde{\alpha}$ which does not tend too fast to infinity and thus defines a convergent solution $y(t)$. This is obviously a statement from analysis. The proof, however, is mostly algebraic and only requires two results from analysis: the implicit function theorem and the Weierstrass division theorem.

To get algebraic power series solutions, say, for a polynomial f , is more delicate.¹ But again, the implicit function theorem and the Weierstrass division theorem hold in this setting (see [Laf1], [Laf2], [Rui, Prop. 5.6]), and this gives the result.

Example. Let \mathbb{K} have characteristic 0. Consider the equation

$$y(t)^3 - t \cdot z(t)^2 = 0$$

in unknown series $y(t) = \sum_{i=0}^\infty y_i t^i$ and $z(t) = \sum_{i=0}^\infty z_i t^i$. We have

$$\begin{aligned} F_0 &= y_0, \\ F_1 &= 3y_0^2 y_1 - z_0^2, \\ F_2 &= 3y_0 y_1^2 + 3y_0^2 y_2 - 2z_0 z_1, \\ F_3 &= y_1^3 + 3y_1 y_2 - 2z_0 z_2 - z_1^2, \end{aligned}$$

and so on. From $F_0 = F_1 = 0$, it follows that $y_0 = 0$ and $z_0 = 0$. Therefore, F_2 is identically 0. We may take $y_1 = 1$ and $z_1 = 1$. This is a permissible choice, though not unique. We could also take $y_1 = 99$ and $z_1 = 501$. But as long as $y_1 \neq 0$, the equality $F_3 = 0$ will prescribe y_2 completely as $y_2 = \frac{1}{3y_1}(z_1^2 - y_1^3)$, whereas z_2 is free. Continuing like this, one easily sees that and how divergent solutions may arise.

The example is so simple that we may also solve the equation $y(t)^3 - t \cdot z(t)^2 = 0$ directly: Write $z(t) = t + \sum \widetilde{z}(t)$ with $\widetilde{z}(t)$ of order ≥ 2 . Then $t \cdot z(t)^2 = t^3 + 2t^2 \cdot \widetilde{z}(t) + \dots$ has order 3 as a series in t , so it admits three cubic roots $\sqrt[3]{t \cdot z(t)^2}$ inside $\mathbb{K}[[t]]$. Any of these roots can be taken for $y(t)$ and can give a solution $(y(t), z(t))$. And if $z(t)$ was formal, respectively convergent, $y(t)$ will be also.

For more complicated examples, such shortcuts of direct computation of the solutions are no longer possible. Nevertheless, general statements about the existence and quality of the solutions can still be made. This will be the content of the next section.

2. THE VARIOUS APPROXIMATION THEOREMS

General references on Artin approximation are, among others, [Te], [Ra1], [PfPo], [Po4], [KMPPR].

We will use two sets of variables, $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$, and power series vectors $f(x, y)$ in both of them, with coefficients in a field \mathbb{K} . The rings of formal, convergent, and algebraic power series in x will be denoted by $\mathbb{K}[[x]]$, $\mathbb{K}\{x\}$, and $\mathbb{K}\langle x \rangle$. They will be equipped with the *Krull topology* for which a system of neighborhoods of 0 is given by the powers $(x)^k$ of the maximal ideal

¹A power series $y(x)$ is called *algebraic* if it satisfies a univariate polynomial equation with polynomial coefficients in x ; see section 6 for more detail.

$(x) = (x_1, \dots, x_n)$. We say that two series $y(x)$ and $z(x)$ agree up to degree c if $y(x) - z(x) \in (x)^c$.

Whenever we have a power series vector $y(x) = (y_1(x), \dots, y_m(x))$ vanishing at 0, we may substitute in $f(x, y)$ the y -variables by $y(x)$ and get a power series vector $f(x, y(x))$ in the x -variables. The question is, When do we get 0 or any other prescribed series $h(x)$?

To simplify the exposition, we assume throughout that the ground field is algebraically closed of characteristic 0. Most results hold for quite arbitrary fields. Whenever we talk about convergent power series, the field \mathbb{K} is assumed to be equipped with an absolute value (multiplicative norm) $|\cdot| : \mathbb{K} \rightarrow \mathbb{R}$ as, e.g., \mathbb{R} , \mathbb{C} , or \mathbb{Q}_p . Our valued fields are not to be confused with the valued fields in valuation theory (fields equipped with a valuation).

We start with the classical implicit function theorem from analysis. It is fundamental for most of the proofs related to Artin approximation.

Implicit Function Theorem. *Let f be a vector of formal, convergent, or algebraic power series in two sets of variables x and y . Assume that $f(0, 0) = 0$, that the number of components of f equals the number of y -variables, and that the relative Jacobian matrix $\partial_y f$ of f with respect to y has evaluation $\partial_y f(0, 0)$ at $(0, 0)$, which is an invertible matrix. There exists a unique vector of formal, convergent, or algebraic power series $y(x)$ with $y(0) = 0$ such that $f(x, y(x)) = 0$.*

Proof. Replacing $f(x, y)$ by the pair $h(x, y) = (x, f(x, y))$ we may assume that $\partial_{xy} h(0, 0)$ is invertible. The solution $y(x)$ of $f(x, y) = 0$ is given by $y(x) = g_2(x, 0)$, where $g(x, y) = (g_1(x, y), g_2(x, y))$ is the inverse to $h(x, y)$ at $(0, 0)$. This is the setting of the inverse function theorem. After composition of h with a linear isomorphism, we may assume that $\partial_{xy} h(0, 0)$ is the identity matrix, so that $h = \text{Id} + q$ where q is a power series vector which is at least quadratic in the variables. Then g is constructed by approximation. Define $g_0 = \text{Id}$, and set

$$g_{k+1} = \text{Id} - g_k \circ q.$$

It is easy to see that the sequence g_k converges in the Krull topology to a vector g of formal power series satisfying $h \circ g = g \circ h = \text{Id}$. This gives the formal inverse to h . For the convergent case, one needs to control the size of the coefficients of the power series expansions of the vectors g_k ; we refer to [HM1, Thm. 6.2, p. 111] for detail. For the algebraic case, see [LT]. □

Here is Artin’s first theorem on approximation as it appeared in his 1968 paper:

Analytic Approximation Theorem (Artin [Ar1, Thm. 1.2]). *Let \mathbb{K} be a valued field of characteristic 0, and let $f(x, y)$ be a vector of convergent power series in two sets of variables x and y . Assume given a formal power series solution $\widehat{y}(x)$ vanishing at 0,*

$$f(x, \widehat{y}(x)) = 0.$$

Then there exists, for any $c \in \mathbb{N}$, a convergent power series solution $y(x)$,

$$f(x, y(x)) = 0,$$

which coincides with $\widehat{y}(x)$ up to degree c ,

$$y(x) \equiv \widehat{y}(x) \text{ modulo } (x)^c.$$

In view of the congruence, one also says that the convergent solutions are dense with respect to the Krull topology in the space of formal solutions. The statement of the theorem was conjectured by S. Lang in 1954 in the complex case $\mathbb{K} = \mathbb{C}$ (cf. [Lan], last paragraph of p. 372). Special cases have been proven, among others, by Greenberg and Spallek for the case of one variable $n = 1$, and by Hironaka and Rossi for equations resulting from isomorphisms between germs of analytic varieties [Gre], [Spa], [HiRo]. Grauert gave a different version of the approximation theorem [Gra], [dJP, p. 291].

An interesting instance is the case where the formal solution is unique. Then, if the system of equations is analytic, the formal solution must already be, by the theorem, convergent.

In the case of linear equations, $\sum_{i=1}^m h_i(x) \cdot y_i = h_0(x)$, for given convergent series $h_i(x)$, the statement of the theorem is equivalent to the faithful flatness of the formal power series ring $\mathbb{K}[[x]]$ over the convergent power series ring $\mathbb{K}\{x\}$. For homogeneous equations $h_0 = 0$, the flatness ensures that the ideal of formal linear relations between h_1, \dots, h_m is generated by the convergent relations. The nonhomogeneous case is then covered by the faithfulness [Bou, Prop. 9, p. 33]. A simple proof for the linear case goes as follows (cf. [Rui, p. 118]): Choose $e > c$, and decompose the given formal solution $\widehat{y}(x)$ into

$$\widehat{y}(x) = y^0(x) + y^1(x) + y^2(x),$$

where $y^0(x)$ is the truncation of $\widehat{y}(x)$ at degree $c - 1$, $y^1(x)$ is the truncation of $\widehat{y}(x) - y^0(x)$ at degree $e - 1$, and $y^2(x)$ is the remaining series of order $\geq e$. Substituting into the linear equation yields

$$\sum_{i=1}^m h_i(x) \cdot [y_i^0(x) + y_i^1(x)] - h_0(x) \equiv 0 \text{ modulo } (x)^e \cdot \mathbb{K}[[x]].$$

But the left-hand side is convergent, so that it actually belongs to $(x)^e \cdot \mathbb{K}\{x\}$. Therefore, since $y^1(x)$ has order $\geq c$, we have

$$\sum_{i=1}^m h_i(x) \cdot y_i^0(x) \in (x)^c \cdot (h_1(x), \dots, h_m(x)) + (x)^e \cdot \mathbb{K}\{x\}.$$

Krull's intersection theorem [AMD, 10.19] now implies that $\sum_{i=1}^m h_i(x) \cdot y_i^0(x)$ already belongs to the ideal $(x)^c \cdot (h_1(x), \dots, h_m(x))$ inside $\mathbb{K}\{x\}$. Writing the sum as a linear combination of $h_1(x), \dots, h_m(x)$ with coefficients $\widetilde{y}_i(x)$ in $(x)^c \cdot \mathbb{K}\{x\}$ provides the convergent solution $y(x) = y^0(x) - \widetilde{y}(x)$ as required.

Algebraic Approximation Theorem (Artin [Ar2, Thm. 1.10]). *Let \mathbb{K} be an arbitrary field, and let $f(x, y)$ be a vector of polynomials or algebraic power series in two sets of variables x and y . Assume given a formal power series solution $\widehat{y}(x)$ vanishing at 0,*

$$f(x, \widehat{y}(x)) = 0.$$

Then there exists, for any $c \in \mathbb{N}$, an algebraic power series solution $y(x)$,

$$f(x, y(x)) = 0,$$

which coincides with $\widehat{y}(x)$ up to degree c ,

$$y(x) \equiv \widehat{y}(x) \text{ modulo } (x)^c.$$

In the two approximation theorems above one may omit the requirement $y(x) \equiv \widehat{y}(x)$ modulo $(x)^c$ and still get an equivalent statement: Indeed, write $\widehat{y}(x) = y^0(x) + \widehat{z}(x)$, where $y^0(x)$ is the truncation of $\widehat{y}(x)$ at degree $c-1$ and where $\widehat{z}(x)$ is a power series vector of order $\geq c$. Write $\widehat{z}(x)$ as $\widehat{z}(x) = \sum_{\alpha} \widehat{v}_{\alpha}(x) \cdot x^{\alpha}$, where $\alpha \in \mathbb{N}^n$ runs over the n -tuples of total degree $\alpha_1 + \dots + \alpha_n = c$, with formal series $\widehat{v}_{\alpha}(x)$. Substituting in $f(x, y)$ the variables y by $y^0(x) + \sum v_{\alpha} \cdot x^{\alpha}$ gives a new system of equations $f'(x, v) = 0$ for variables $v = (v_{\alpha})_{\alpha}$. Clearly, $\widehat{v}(x)$ is a formal solution of $f'(x, v) = 0$. Then the existence of a convergent or algebraic solution $v(x)$ of $f'(x, v) = 0$ (without any approximation property with respect to $\widehat{v}(x)$) yields the required convergent or algebraic solution $y(x)$ of $f(x, y) = 0$ satisfying $y(x) \equiv \widehat{y}(x)$ modulo $(x)^c$.

Nested Algebraic Approximation Theorem (Popescu, Spivakovsky, [Po1], [Po2], [Rot], [Spi2], [Og], [Te], [Sw], [Qu], [BDLv], [CPR]). *Let \mathbb{K} be an arbitrary field, and let $f(x, y)$ be a vector of polynomials or algebraic power series in two sets of variables x and y . Assume given a formal power series solution $\widehat{y}(x)$ vanishing at 0,*

$$f(x, \widehat{y}(x)) = 0$$

so that each component $\widehat{y}_i(x)$ of $\widehat{y}(x)$ only depends on the first n_i variables x_1, \dots, x_{n_i} , for some $1 \leq n_i \leq n$. Then there exists, for any $c \in \mathbb{N}$, an algebraic power series solution $y(x)$,

$$f(x, y(x)) = 0$$

which coincides with $\widehat{y}(x)$ up to degree c ,

$$y(x) \equiv \widehat{y}(x) \text{ modulo } (x)^c,$$

and for which $y_i(x)$ again only depends on the first n_i variables x_1, \dots, x_{n_i} .

This theorem is much harder than all the other theorems in this text, and it will not be proven here. It follows from Popescu's General Néron Desingularization Theorem [Po1], [Po2], [Spi2], [Sw], [Te]: *Any regular ring morphism between Noetherian rings is the inductive limit of smooth morphisms.* A direct elementary proof of nested approximation seems to be unknown. Nested approximation does not hold if the power series vector f is just analytic but not algebraic. A counterexample to this was given by Gabrielov [Ga1], [Ga2]; see section 7 on formal and analytic relations.

In a similar vein, Becker gave an example of a polynomial equation which admits a formal solution whose components depend on *disjoint* sets of variables and which does not admit a convergent solution with the same property [Be].

Parametrized Approximation Theorem (Płoski [P11], [P12], [P13]). *Let \mathbb{K} be a valued field of characteristic 0, and let $f(x, y)$ be a vector of convergent power series in two sets of variables x and y . Assume given a formal power series solution $\widehat{y}(x)$ vanishing at 0,*

$$f(x, \widehat{y}(x)) = 0.$$

Then there exists a convergent power series solution $y(x, z)$ depending on x_1, \dots, x_n and new variables $z = (z_1, \dots, z_s)$,

$$f(x, y(x, z)) = 0,$$

and a vector of formal power series $\widehat{z}(x) = (\widehat{z}_1(x), \dots, \widehat{z}_s(x))$ vanishing at 0 such that

$$\widehat{y}(x) = y(x, \widehat{z}(x)).$$

This statement can be understood as follows. The vector $y(x, z)$ induces, for any choice of $z(x) \in \mathbb{K}[[x]]^s$ with $z(0) = 0$, a solution $y(x, z(x))$ of $f(x, y) = 0$. This gives a map $\widehat{\alpha} : (x) \cdot \mathbb{K}[[x]]^s \rightarrow \widehat{\mathcal{Y}}(f)$, where $\widehat{\mathcal{Y}}(f) \subset (x) \cdot \mathbb{K}[[x]]^m$ denotes the set of all formal solutions of $f(x, y) = 0$. It maps the vector $\widehat{z}(x)$ to the given formal solution $\widehat{y}(x)$. As $y(x, z)$ is convergent, $\widehat{\alpha}$ induces a map $\alpha : (x) \cdot \mathbb{K}\{x\}^s \rightarrow \mathcal{Y}(f) = \widehat{\mathcal{Y}}(f) \cap \mathbb{K}\{x\}^m$ into the set of convergent solutions.

The solution sets $\widehat{\mathcal{Y}}(f)$ and $\mathcal{Y}(f)$ are hence only *partially* parametrized, in the first case nearby the formal solution $\widehat{y}(x)$. A more detailed description of these sets in the case of one x -variable is provided in [HW].

Strong Approximation Theorem I. *Let $f(x, y)$ be a vector of formal power series in two sets of variables x and y . Assume given a sequence $\{y^{(k)}(x)\}_{k \geq 1}$ of approximate solutions of $f(x, y) = 0$ up to degree k and vanishing at 0, i.e., such that*

$$f(x, y^{(k)}(x)) \in (x)^k \cdot \mathbb{K}[[x]]^m \text{ for all } k \geq 1.$$

Then there exists an exact formal solution $\widehat{y}(x)$ of $f(x, y) = 0$, say

$$f(x, \widehat{y}(x)) = 0.$$

For a similar statement, see Grauert's approximation theorem [Gra], [dJP, Thm. 8.2.2].

Strong Approximation Theorem II. *Let $f(x, y)$ be a vector of formal power series in two sets of variables x and y . For any $c \in \mathbb{N}$, there exists an integer $e \in \mathbb{N}$ such that if $f(x, y) = 0$ admits an approximate solution $\overline{y}(x)$ up to degree e and vanishing at 0,*

$$f(x, \overline{y}(x)) \in (x)^e \cdot \mathbb{K}[[x]]^m,$$

then there exists an exact formal solution $\widehat{y}(x)$ to $f(x, y) = 0$, say

$$f(x, \widehat{y}(x)) = 0,$$

and such that $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^c$.

In concrete applications, the two statements are of equal use, since the existence of the bound e of version II in dependence of c is most often only of theoretical nature, so in any case one has to find approximate solutions up to arbitrary degree as in version I to ensure the existence of an exact formal solution. We give an application of strong approximation in a later section.

Notice also that in the assumption of the Strong Approximation Theorem I it is not required that the approximate solutions $y^{(k)}(x)$ coincide with each other up to a certain degree, i.e., that, for $k' \geq k$, the truncation of $y^{(k')}(x)$ at degree k need not equal $y^{(k)}(x)$. Otherwise, if we always had $y^{(k')}(x) \equiv y^{(k)}(x)$ modulo $(x)^k$, the limit $\widehat{y}(x) = \lim_k y^{(k)}(x)$ would trivially exist in the Krull topology and give the exact formal solution. So it is this more general assumption which makes the result very useful.

For linear equations, the value of e can be determined by the Weierstrass division theorem for ideals [Gra], [Hir]. Indeed, consider an inhomogeneous equation $h_0(x) = \sum_{i=1}^m h_i(x) \cdot y_i$ with given formal power series $h_0(x)$ and $h_1(x), \dots, h_m(x)$.

There exists a solution $\widehat{y}(x)$ if and only if $h_0(x)$ belongs to the ideal I generated by h_1, \dots, h_m . This is equivalent to saying that the remainder of the division of h_0 by I is zero. If $h_0 \notin I$, the remainder is nonzero and has a certain order o . Hence there exists no approximate solution $\overline{y}(x)$ up to degree $e = o + 1$. If $h_0 \in I$, the exact solution $\widehat{y}(x)$ exists and there is nothing to show.

Uniform Strong Approximation Theorem. *Let $n, m, \ell \in \mathbb{N}$ be integers, and fix variables $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$. There exists a function $\beta = \beta_{n,m,\ell} : \mathbb{N} \rightarrow \mathbb{N}$ depending only on n, m , and ℓ such that the following holds. For all polynomial vectors $f(x, y)$ of total degree $\leq \ell$ in x and y , for all $c \in \mathbb{N}$, and for all approximate solutions $\overline{y}(x)$ of $f(x, y) = 0$ up to degree $\beta(c)$, say*

$$f(x, \overline{y}(x)) \in (x)^{\beta(c)},$$

there exists an exact algebraic solution $y(x)$ to $f(x, y) = 0$,

$$f(x, y(x)) = 0,$$

such that $y(x) \equiv \overline{y}(x)$ modulo $(x)^c$.

Remarks. Greenberg has shown that in the one-variable case $n = 1$, one can choose for β a linear function $\beta(c) = ac + b$ in c , for suitable integers $a, b \in \mathbb{N}$ [Gre]. In the general case, the smallest function β satisfying the conclusion of the theorem is called the *Artin function* [Hic], [Spi1], [Ron1], [Ron2].

The three strong approximation theorems can be combined with the convergent, respectively algebraic, approximation theorem to ensure from a sufficiently good approximate solution the existence of a convergent, respectively algebraic, solution. Artin himself mentions that he is not aware of applications where the first two strong approximation theorems are not sufficient and where the uniform strong version is really needed [Ar2, bottom of p. 51].

For one x -variable there is an analogous strong approximation theorem by Denef and Lipshitz for ordinary differential equations with polynomial coefficients, but which omits the density requirement $y(x) \equiv \overline{y}(x)$ modulo $(x)^c$ [DL2, Thm. 2.10 and Thm. 2.14]. Their argument is built on ultraproducts. A sufficient assumption is that the ground field is real closed as \mathbb{R} or algebraically closed as \mathbb{C} . The result does not hold for arbitrary fields. The ultraproduct technique also provides an alternative proof for the original strong approximation theorem on the approximate vs. formal solutions of formal power series equations.

Example (Denef and Lipshitz, [DL2, remarks 2.12 and 2.15]). Take for \mathbb{K} the ring of rational functions $\mathbb{K} = \mathbb{R}(t)$ over \mathbb{R} and consider the system of ordinary differential equations

$$xy' - (z + x)y - 1 = 0, \quad z' = 0,$$

in unknown series $y(x)$ and $z(x)$. Clearly, $z(x)$ must be a constant $\alpha \in \mathbb{K}$. We distinguish two cases: If $\alpha \in \mathbb{K} \setminus \mathbb{N}$, the system has the solution

$$y(x) = \sum_{i=1}^{\infty} (-1)^i \frac{x^{i-1}}{\alpha^{\dot{i}}},$$

with $\alpha^{\dot{i}} = \alpha(\alpha - 1) \cdots (\alpha - i + 1)$ the falling factorial. If $\alpha \in \mathbb{N}$, there is no exact solution. Denef has shown that there exists a polynomial $P \in \mathbb{K}[w_1, \dots, w_s]$ so that an element $\alpha \in \mathbb{K}$ belongs to \mathbb{N} if and only if there are $w_1, \dots, w_s \in \mathbb{K}$ so that

$P(\alpha, w_1, \dots, w_s) = 0$ (i.e., \mathbb{N} is *diophantine* over $\mathbb{R}(t)$) [Den]. Adding this equation to the two differential equations, one gets a system which admits, for every c , an approximated solution $(\bar{y}(x), \bar{z}(x))$ up to degree c , with $\bar{z}(x) = \alpha \in \mathbb{N}$. But, by the preceding observation, the system cannot have an exact solution. So strong approximation does not hold for $\mathbb{R}(t)$.

In the case of several variables, the situation for systems of partial differential equations splits. Even for the real case, strong approximation fails [DL2, Cor. 4.10]. Over an algebraically closed field and waiving the density condition, one has again a strong approximation result, even in a more general setting than differential equations. The proof we give is of a different flavor and was communicated to us by S. Mori. Call a map $\mathbb{K}[[x]]^m \rightarrow \mathbb{K}[[z]]^q$ between power series spaces *textile* if the coefficients of the image series are *polynomials* in the coefficients of the input series. See [HM1], [BH], [HW] for a detailed discussion of such maps.

Strong Approximation Theorem for Textile Maps. *Assume that \mathbb{K} is an uncountable algebraically closed field, and let x denote again a vector of variables. Let $G : \mathbb{K}[[x]]^m \rightarrow \mathbb{K}[[z]]^q$ be a textile map. There exists an $\ell \in \mathbb{N}$ depending on G such that, if $G = 0$ admits an approximate solution $\bar{y}(x) \in \mathbb{K}[[x]]^m$ up to degree ℓ ,*

$$G(\bar{y}(x)) \equiv 0 \pmod{(x)^\ell},$$

then there exists an exact solution $y(x) \in \mathbb{K}[[x]]^m$,

$$G(y(x)) = 0.$$

Remark. For the case of partial differential equations with polynomial coefficients over \mathbb{C} (which induce textile maps), this result appears in [DL2, Thm. 2.10 and Rem. 2.11]. It is in general not possible to find a solution $y(x)$ which coincides with $\bar{y}(x)$ up to a prescribed degree, just take the two differential equations from the example before, but now over \mathbb{C} [DL2, Rem. 2.12 and 2.15]. It seems to be an open problem to find sufficient conditions on G beyond the case where $G(y(x))$ is given by substitution of y by $y(x)$ in a power series vector $f(x, y)$ which ensures that the exact solution can be chosen to coincide with the approximated solution up to a prescribed degree.

3. PROOF OF THE ANALYTIC AND ALGEBRAIC APPROXIMATION THEOREMS

We shall provide a proof that applies, with minor modifications, also to the other approximation theorems. The necessary changes are described in the next section. Artin's original proof of the analytic case in [Ar1] is somewhat shorter, though more computational. All ideas and constructions presented below already appear in Artin's papers [Ar1], [Ar2], complemented by an additional idea due to Płoski [P11].

The proof goes by induction on the number n of x -variables. So let $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_m)$, and $f = (f_1, \dots, f_r)$. If $n = 0$, the series $\hat{y}(x)$ is a vector of constants in the field \mathbb{K} and there is nothing to prove. So let $n \geq 1$.

Here is the strategy. We will construct from f (and using also a small amount of information about the formal solution $\hat{y}(x)$) a new vector of convergent power series $f' = f'(x', w)$ in the first $n - 1$ components $x' = (x_1, \dots, x_{n-1})$ of the x -variables and in new variables $w = (w_0, \dots, w_{d-1})$ such that the existence of a formal or convergent solution $y(x)$ to $f(x, y) = 0$ is *equivalent* to the existence of a formal or convergent solution $w(x')$ of $f'(x', w) = 0$. We will show that the

formal solutions $\widehat{y}(x)$ of the first system are in one-to-one correspondence with the formal solutions $\widehat{w}(x')$ of the latter system, and that the same correspondence holds for the convergent solutions $\widetilde{y}(x)$ and $\widetilde{w}(x')$. Therefore, induction applies: by the existence of the formal solution $\widehat{w}(x')$ and the induction hypothesis, the system $f'(x', w) = 0$ admits a convergent solution $\widetilde{w}(x')$, and going backwards we get the required convergent solution $\widetilde{y}(x)$ of the original system $f(x, y(x)) = 0$.

Step 1. Reduction to the case where an $(r \times r)$ -minor g of the relative Jacobian matrix $\partial_y f$ of f with respect to y does not vanish at $\widehat{y}(x)$, say $g(x, \widehat{y}(x)) \neq 0$ (recall that r is the number components of the vector f). This step is necessary to be able to apply the Jacobian criterion of regularity.

Let I be the ideal of $\mathbb{K}\{x, y\}$ generated by f_1, \dots, f_r . We may clearly enlarge I and replace it by the ideal of all series vanishing at $\widehat{y}(x)$. This will allow us to assume that I is a prime ideal. We may consider this ideal as the *saturation* of the ideal generated by the original series f_1, \dots, f_r with respect to the formal solution $\widehat{y}(x)$.

Let s be the height of I , i.e., the codimension of the germ of the analytic variety X defined by I in $(\mathbb{K}^{n+m}, 0)$. It is well known that X equals an irreducible component of the germ of a complete intersection variety X^* in $(\mathbb{K}^{n+m}, 0)$; see, e.g., [Mu, Cor. 4, p. 44]. Therefore there exists an ideal $I^* \subset I$ generated by a regular sequence $f^* = (f_1, \dots, f_s)$ of length s of elements of I such that I is a prime component of I^* . Denoting by J the intersection of the other prime components of I^* , we have $I^* = I \cap J$.

By the Jacobian criterion [Mu, Cor. 1 to Prop. 2, p. 168], there exists an $(r \times r)$ -minor g of the Jacobian matrix $\partial_{xy} f^*$ which does not belong to I^* . Hence $g(x, \widehat{y}(x)) \neq 0$ for our given formal solution $\widehat{y}(x)$. Taking the partial derivatives of $f^*(x, \widehat{y}(x)) = 0$ with respect to x_1, \dots, x_n we get

$$\partial_{x_i} f^*(x, \widehat{y}(x)) = - \sum_{k=1}^r \partial_{x_i} \widehat{y}_k(x) \cdot \partial_{y_k} f^*(x, \widehat{y}(x)).$$

The equality shows that the minor g can actually be chosen within the relative Jacobian matrix $\partial_y f^*$ of f^* with respect to the y -variables.

Assume now that we have a convergent or algebraic solution $y(x)$ of $f^*(x, y) = 0$ approximating the given formal solution $\widehat{y}(x)$ up to a prescribed degree. We claim that then $y(x)$ is also a solution of $f(x, y) = 0$ for all $f \in I$. But $g(x, y(x)) \neq 0$ because $g(x, \widehat{y}(x)) \neq 0$ and because $y(x)$ approximates $\widehat{y}(x)$ to high order. Moreover g vanishes on the singular locus of X and hence belongs, by Rückert's Nullstellensatz [Rui], to the radical of the ideal $I + J$ defining the intersection of the component of X defined by I with the union of the other components, say $g^b \in I + J$ for some $b \in \mathbb{N}$. It follows that $g^b \cdot I \subset I \cdot J \subset I \cap J = I^*$. Hence $g^b(x, y(x)) \cdot f(x, y(x)) = 0$ for all $f \in I$. From $g(x, y(x)) \neq 0$, we conclude that actually $f(x, y(x)) = 0$ for all $f \in I$. Therefore $f(x, y(x)) = 0$ for all $f \in I$ is equivalent to $f(x, y(x)) = 0$ for all $f \in I^*$. Consequently, we may replace I by I^* and assume henceforth that the number r of generators f_1, \dots, f_r of I equals the height of I . This establishes the first reduction step.

Step 2. Let $y(x)$ be a vector of formal, convergent, or algebraic power series such that $g(x, y(x))$ is nonzero of order d . We show for all such vectors that the system

of equations $f(x, y(x)) = 0$ is equivalent to a system of ideal memberships

$$h_j(x, z(x)) \in (g(x, z(x)))^2$$

for a suitable vector $h = (h_1, \dots, h_r)$ of convergent, respectively algebraic, power series in x_1, \dots, x_n and $z = (z_1, \dots, z_m)$, and where $z(x)$ is a vector of formal, convergent, or algebraic power series which are *polynomial* in the last variable x_n of degree $< d$. Here, $(g(x, z(x)))^2$ denotes the ideal in $\mathbb{K}[[x]]$, respectively $\mathbb{K}\{x\}$, generated by $g(x, z(x))^2$, according to the setting. Therefore, $z(x)$ is determined by its finitely many coefficient vectors in front of x_n^j , which are now power series vectors in the $n - 1$ variables $x' = (x_1, \dots, x_{n-1})$. The vector h is constructed explicitly from f and g . The proof of the equivalence is the same for formal, convergent, and algebraic series.

After a linear generic coordinate change in the x -variables, it can be assumed that $g(x, y(x))$ is x_n -regular of order d .² The idea then is to divide the components $y_i(x)$ of $y(x)$ by $g(x, y(x))$, respectively its square $g(x, y(x))^2$, using Weierstrass division

$$y_i(x) = v_i(x) + z_i(x),$$

where $v_i(x)$ is a power series in the ideal generated by $g(x, y(x))$, respectively $g(x, y(x))^2$, and where $z_i(x)$ is a polynomial in x_n of degree $\leq d - 1$, respectively $\leq 2d - 1$, with coefficients power series in $x' = (x_1, \dots, x_{n-1})$. The distinction between $g(x, y(x))$ and $g(x, y(x))^2$ was suggested by Płoski and is well suited for technical reasons appearing later in the proof: Split the y -variables into two groups y' and y'' , where the first indicates the columns of $\partial_y f$ used for the minor g , and y'' the remaining components. We may assume, after renumeration of the components of y , that $y = (y', y'')$. Then all components of $y'(x)$ will be divided by $g(x, y(x))$, and those from $y''(x)$ by $g(x, y(x))^2$. Denote by $v(x)$ and $z(x)$ the resulting vectors, $y(x) = v(x) + z(x)$, with components $v_i(x)$ and $z_i(x)$.

Lemma. *The series $g(x, y(x))$ and $g(x, z(x))$ generate the same ideals of $\mathbb{K}[[x]]$, $\mathbb{K}\{x\}$, or $\mathbb{K}\langle x \rangle$, according to the setting.*

Proof. Taylor expansion of $g(x, z(x)) = g(x, y(x) - v(x))$ shows that all resulting terms are multiples of $g(x, y(x))$, the first one being $g(x, y(x))$ itself. Factoring $g(x, y(x))$ from these terms produces a unit $u(x)$ as the quotient, $g(x, z(x)) = g(x, y(x)) \cdot u(x)$, and this shows the claim.

Let us from now on write y and z for $y(x)$ and $z(x)$. To further ease the notation, we will also often omit in expressions like $f(x, z)$ the reference to the x -variable and simply write $f(z)$. By the lemma, we have $v = a \cdot g(z)$ with some vector $a = (a', a'') \in \mathbb{K}[[x]]^m$ whose latter components a'' are multiples of $g(z)$ (here, a'' collects the same group of components of a as y'' did for y). The partial derivatives $(\partial_{y_1} f_j, \dots, \partial_{y_m} f_j)$ of the components f_j of f are taken as a row vector, while f is written as a column vector.

Substituting the decomposition $y = a \cdot g(z) + z$ in $f(y) = 0$ gives by Taylor expansion the equivalent system of equations,

$$f(z) + g(z) \cdot \partial_y f(z) \cdot a + g(z)^2 \cdot q(a, g(z)) = 0,$$

²A series in x_1, \dots, x_n is x_n -regular of order d if a pure x_n -monomial x_n^d appears in its expansion with a nonzero coefficient.

where $q(w, t) = q(x, w, t) \in \mathbb{K}\{x, w, t\}^r$ is a convergent power series vector in the variables $x = (x_1, \dots, x_n)$, $w = (w_1, \dots, w_m)$ and a single variable t . Note that q is at least quadratic in w . We now use that g is a maximal minor of $\partial_y f$, of size r , and that g is given by the columns of $\partial_y f$ corresponding to the y' -variables. Therefore we can write the diagonal matrix $g \cdot \mathbb{1}_r$ as a product $g \cdot \mathbb{1}_r = \partial_{y'} f \cdot \partial_{y'}^* f$, where $\partial_{y'}^* f \in \mathbb{K}\{x, y\}^{r \times r}$ is the adjoint matrix of the submatrix $\partial_{y'} f$ of $\partial_y f$ defining g . It follows that

$$g(z)^2 \cdot q(a, g(z)) = g(z) \cdot \partial_{y'} f(z) \cdot \partial_{y'}^* f(z) \cdot q(a, g(z)).$$

But recall that the components of a'' are multiples of $g(z)$, say $a'' = \tilde{a}'' \cdot g(z)$ for some \tilde{a}'' . Hence, by the same trick, we also have

$$g(z) \cdot \partial_{y'} f(z) \cdot a'' = g(z) \cdot \partial_{y'} f(z) \cdot \partial_{y'}^* f(z) \cdot \partial_{y'} f(z) \cdot \tilde{a}''.$$

Combining these two formulas gives the equivalent system of equations

$$f(z) + g(z) \cdot \partial_{y'} f(z) \cdot \phi(a, z) = 0$$

for z and a , where

$$\phi(x, w, z) = w' + \partial_{y'}^* f(z) \cdot q(w, g(z)) + \partial_{y'}^* f(z) \cdot \partial_{y''} f(z) \cdot \tilde{w}''$$

with $w = (w', \tilde{w}'')$. Observe that ϕ is a convergent, respectively algebraic, power series vector with r components vanishing at 0 for which $\partial_w \phi(0, 0, 0)$ is an invertible $(r \times r)$ -matrix. By the implicit function theorem there is a unique convergent, respectively algebraic, power series vector $\psi(x, w, z)$ vanishing at 0 such that $\psi(x, \phi(x, w, z), z) = w'$. Setting $b(x, z) = \phi(x, a, z)$, we may therefore rewrite the system from before as

$$f(x, z) + g(x, z) \cdot \partial_{y'} f(x, z) \cdot b(x, z) = 0$$

with a convergent, respectively algebraic, power series vector b defined uniquely by f . Conversely, any choice of z and b satisfying this system produces, going backwards, a solution y of $f(x, y) = 0$.

We now multiply the preceding system from the left with the matrix $\partial_{y'}^* f$. Since $\partial_{y'}^* f(z)$ is not a zero divisor in the ring of $(r \times r)$ -matrices, we get the equivalent system

$$\partial_{y'}^* f(z) \cdot f(z) + g(z) \cdot \partial_{y'}^* f(z) \cdot \partial_{y'} f(z) \cdot b = 0.$$

In this system we can replace $\partial_{y'}^* f \cdot \partial_{y'} f$ by $g \cdot \mathbb{1}_r$ yielding finally

$$\partial_{y'}^* f(z) \cdot f(z) + g(z)^2 \cdot b = 0.$$

Setting $h = (h_1, \dots, h_r) = \partial_{y'}^* f \cdot f$, we get the equivalent ideal membership condition

$$h_j(z) \in (g(z)^2)$$

in $\mathbb{K}[[x]]$, respectively $\mathbb{K}\{x\}$ or $\mathbb{K}\langle x \rangle$, for $j = 1, \dots, r$. This is valid simultaneously for the formal, convergent, and algebraic setting, while h and g are both convergent, respectively algebraic, power series. The equivalence claimed at the beginning of this step is established. This transcription of the original problem to an ideal membership is often referred to as the reduction to the setting of Tougeron's implicit function theorem [Tou].

Note here for later use that if f is a polynomial vector, say of total degree ℓ in x and y , then h is also a polynomial vector whose total degree in x and z is bounded from above by $\ell + (\ell - 1)^{r-1} \leq \ell + (\ell - 1)^{m-1}$.

Step 3. Reduction of the membership $h_j(x, z(x)) \in (g(x, z(x)))^2$ for $j = 1, \dots, r$ to an equivalent system of convergent, respectively algebraic, equations for power series in $n - 1$ variables. Again we work simultaneously with formal, convergent and algebraic power series vectors $z(x)$.

Recall that $z = z(x)$ is a vector of polynomials in x_n of degree $< d$ with coefficients $w_i = w_i(x')$ which are formal, convergent, or algebraic power series vectors in $x' = (x_1, \dots, x_{n-1})$ and that $g(x, z(x))$ is x_n -regular of order d . The passage from the above membership condition for $z(x)$ to an equivalent system of convergent or algebraic equations $f'(x', w) = 0$ for unknown series $w(x')$ in $n - 1$ variables is ensured by the Weierstrass division theorem stated after the proofs of the approximation theorems. Again, going backwards, a solution $w(x')$ of the latter will provide a solution of the former by setting $z(x) = \sum_{i=0}^{d-1} w_i(x') \cdot x_n^i$.

We note here for later use in the proof of the uniform strong approximation theorem that if h and g are polynomial vectors of total degree ℓ , respectively k , in x and z , then f' will also be a polynomial vector since then the division algorithm is finite. Its total degree in x and z will be bounded from above by $\ell + (\ell - d + 1)(k - d) \leq \ell(1 + k)$.

Step 4. The proof of the analytic and algebraic approximation theorems finishes by induction: Starting with our formal solution $\widehat{y}(x)$ of $f(x, y) = 0$, we get, following the preceding constructions, a formal solution $\widehat{w}(x')$ of an equivalent convergent system $f'(x', w) = 0$ in $n - 1$ variables, for some convergent vector $f' \in \mathbb{K}\{x', w\}^{r'}$. By induction on the number of variables there exists, for any $c \in \mathbb{N}$, a vector $\widetilde{w}(x')$ of convergent solutions in $\mathbb{K}\{x'\}^d$ to the system $f'(x', w) = 0$, and such that $\widetilde{w}_i(x') \equiv \widehat{w}_i(x')$ modulo $(x')^c$. It now suffices to go backwards through all equivalences to find the required convergent solution $\widetilde{y}(x)$ of $f(x, y) = 0$ with $\widetilde{y}(x) \equiv \widehat{y}(x)$ modulo $(x)^c$. This proves the analytic and algebraic approximation theorems. \square

4. PROOFS OF THE PARAMETRIZATION THEOREM AND THE FOUR STRONG APPROXIMATION THEOREMS

We now indicate how to modify the proof of the analytic and algebraic approximation theorems to establish the other theorems (aside from the nested approximation theorem which is more difficult to prove).

(a) The proof of the parametrization theorem in both settings, the convergent and the algebraic one, does not need any substantial changes, provided that one refers in the induction step also to the stronger assertion as indicated in the theorem. The partial parametrization of the solution set of $f(x, y) = 0$ nearby a formal solution $\widehat{y}(x)$ is then constructed by working again backwards from the partial parametrization of the system $f'(x', w) = 0$.

(b) The strong approximation theorem I is proven by taking in Step 1 the saturation with respect to all sequences of approximate solutions. To this end, let I be the ideal of convergent power series $h(x, y)$ for which the order of $h(x, y^{(k)}(x))$ tends towards infinity for all sequences $y^{(k)}(x)$ of power series vectors for which the order of $f(x, y^{(k)}(x))$ tends towards infinity. This is a prime ideal and the constructions of Step 1 (on page 604) apply again to I . One may therefore start from the beginning with a vector f whose components generate a prime ideal I and assume, after possibly passing to a subsequence, that $f(x, y^{(k)}(x)) \equiv 0$ modulo $(x)^k$ for all k . Denote by r the height of I . As in Step 1 the ideal I is replaced by an ideal I^* generated by r series f_1, \dots, f_r and such that a minor g of $\partial_{xy}f$ does

not belong to I^* and hence also not to I . Taking partial derivatives with respect to the x -variables of the congruence $f(x, y^{(k)}(x)) \equiv 0$ modulo $(x)^k$, one sees that this minor can actually be chosen from the relative Jacobian matrix $\partial_y f$.

The key point then is to observe that there is an integer d so that $g(x, y^{(k)}(x))$ has the same order d for infinitely many k : By definition of I and since $g \notin I$, there is a subsequence of $y^{(k)}(x)$ for which the order of $g(x, y^{(k)}(x))$ remains bounded. Upon passing possibly to a further subsequence, we may assume that this order is actually constant, say d . From this moment on, the proof is essentially identical to the given one. A linear coordinate change in the x -variables allows us to assume that $g(x, y^{(k)}(x))$ is x_n -regular of order d (this requires that \mathbb{K} is infinite; if \mathbb{K} is finite, one has to use a nonlinear coordinate change, also known as Nagata's trick, see [Mu, p. 2], yielding an order d' of $g(x, y^{(k)}(x))$ in x_n possibly larger than d). Then, the vectors $y^{(k)}(x)$ are divided componentwise by $g(x, y^{(k)}(x))$. In the rest of the proof all equations of type $[\dots] = 0$ for series $y(x)$ have to be replaced by the respective congruences of type $[\dots] \equiv 0$ modulo $(x)^k$ for the sequence $y^{(k)}(x)$. The same modifications have to be applied to the ideal membership conditions. In this way one reduces to a system of equations in one variable less which admits approximate solutions up to any degree. Induction applies to ensure the existence of an exact formal solution of this system, and going backwards one obtains the exact formal solution of the original system.

(c) The strong approximation theorem II is proven as follows. We first reduce again to prime ideals. So let I be the ideal generated by the components f_1, \dots, f_r of f . Choose an irredundant primary decomposition $I = I_1 \cap \dots \cap I_t$, and let $J_i = \sqrt{I_i}$ be the associated prime ideals. As $\mathbb{K}[[x]]$ is Noetherian, there exists an integer u such that $J_i^u \subset I_i$ for all i . Let e_i be the bound associated to J_i by the theorem in case of prime ideals. Then $e = u \cdot (e_1 + \dots + e_t)$ will work for I , cf. [Wa, proof of Lemma 5, p. 133]. Namely,

$$J_1^u \dots J_t^u \subset I_1 \dots I_t \subset I_1 \cap \dots \cap I_t = I,$$

so that any approximate solution $\overline{y}(x)$ for I up to degree e is also an approximate solution of some I_i up to degree e_i , for some i . By assumption, I_i then admits an exact formal solution $\widehat{y}(x)$ with $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^e$. From $I \subset I_i$ it follows that $\widehat{y}(x)$ is also an exact formal solution for I .

So we may assume that I is prime. Let s be its height, and choose again an $(s \times s)$ -minor g of the Jacobian matrix so that $g \notin I$. We will now determine a suitable value for e . So fix some e , and consider an approximate solution $\overline{y}(x)$ of $f(x, y) = 0$ up to degree e . If $g(x, \overline{y}(x)) \equiv 0$ modulo $(x)^e$, increase e by 1 and repeat. If this continues indefinitely, replace I by the ideal $I' = I + (g)$. It has height $> s$ and satisfies again the assumption of the theorem. So we can apply induction on the height of I . Otherwise, if $g(x, \overline{y}(x)) \not\equiv 0$ modulo $(x)^e$ for some e , we may apply a generic linear triangular coordinate change to turn $g(x, \overline{y}(x))$ into an x_n -regular series of order e . Now we are in the situation where the proof of the analytic approximation theorem applies and can be repeated with the modifications as indicated in part (b) above. This proves version II of the strong approximation theorem. □

(d) The proof of the uniform strong approximation theorem requires more effort. We follow Artin's original line of arguments in [Ar2]. For a model-theoretic proof using ultraproducts, see [BDLv]. Let us say that a polynomial ideal has degree ℓ if it can be generated by polynomials of degree $\leq \ell$ but not of degree $< \ell$. Also, we

call $y(x)$ a solution for the ideal I if $h(x, y(x)) = 0$ holds for all $h \in I$, and similarly for approximated solutions.

Step 1. Reduction to prime ideals. Assume that the theorem holds whenever the components of f generate a prime ideal. Let β' denote the associated function for prime ideals as described in the theorem. We will construct from these data a function β which works for all ideals.

Let I be the ideal generated by the components of the given vector f . This is not necessarily a prime ideal nor a radical, so let $I = I_1 \cap \dots \cap I_t$ be an irredundant primary decomposition. By the work of Hermann [Her], the number t and the degrees ℓ_j of the primary components I_j and associated primes $J_j = \sqrt{I_j}$ of I are bounded by a function $\lambda : \mathbb{N} \rightarrow \mathbb{N}$ depending only on n, m , and the degree ℓ of I . Moreover, $J_j^u \subset I_j$ for some $u \leq \lambda$ and all j .

Now argue as follows: Let $b' = \beta'_{n,m,\ell'}(c)$ be the value of the function β' at c , with respect to the parameter values n, m , and $\ell' = \lambda(n, m, \ell)$. If $\bar{y}(x)$ is an approximate solution for I up to degree $u \cdot t \cdot b'$, then, because of the inclusion

$$I_1 \cdots I_t \subset I_1 \cap \dots \cap I_t = I,$$

$\bar{y}(x)$ is also an approximate solution for each I_j up to degree $u \cdot b'$, and consequently for each J_j up to degree b' . Pick one such J_j . Assuming the result to be true for prime ideals with respect to the function β' , we get an exact formal solution $\hat{y}(x)$ for J_j . But $I \subset J_j$, so $\hat{y}(x)$ is also a solution for I .

In view of these considerations, we are led to define β as $\beta = \lambda^2 \cdot \beta'$. This function will then work for all ideals, establishing the claimed reduction to the case of prime ideals.

Step 2. Reduction to complete intersection ideals. This goes as follows. Fix an integer $s \leq n$. Assume by induction that the theorem holds for *all* ideals of height $> s$ (not just the prime ideals). Assume also that it holds for complete intersection ideals, i.e., ideals generated by a regular sequence. Let β' denote the associated function for these two cases as described in the theorem. We will construct from these data a function β which works for all prime ideals of height s .

Now let I be a prime ideal, and let s be its height. Let $f = (f_1, \dots, f_r)$ generate I , for some $r \geq s$. The height s equals the codimension of the irreducible subvariety X of $\mathbb{A}_{\mathbb{K}}^n \times \mathbb{A}_{\mathbb{K}}^m$ defined by f . Among the generators f_1, \dots, f_r one can choose s polynomials which form a regular sequence, i.e., so that the subvariety X^* defined by them in $\mathbb{A}_{\mathbb{K}}^n \times \mathbb{A}_{\mathbb{K}}^m$ is a complete intersection (which means that X^* can be defined by as many equations as the codimension indicates, and X is an irreducible component of X^*); see again [Mu, Cor. 4, p. 44].

We may of course assume that the first s polynomials f_1, \dots, f_s are these elements. Denote by I^* the ideal generated by them. In the prime decomposition of I^* , the ideal I appears as a prime component. Denote by J the intersection of the other ideals, so that $I^* = I \cap J$, and set $Y = V(J) \subset \mathbb{A}_{\mathbb{K}}^n \times \mathbb{A}_{\mathbb{K}}^m$. As I is prime of height s , the Jacobian criterion [Mu, Cor. 1 to Prop 2, p. 168] ensures that there exists an $(s \times s)$ -minor g of the Jacobian matrix $\partial_{xy} f$ of $f = (f_1, \dots, f_r)$ such that $g \notin I$. Choose such a minor, and let $\ell' = \max\{\ell, (\ell - 1)^s\}$. Assume given an approximate solution $\bar{y}(x)$ for I up to degree $\beta'(c) = \beta'_{n,m,\ell'}(c)$. We distinguish two cases.

(i) If $\bar{y}(x)$ is also an approximate solution of $g(x, y) = 0$ up to degree $\beta'(c)$, we may replace I by the ideal I' generated by I and g . This ideal has height

$s + 1$ since I is prime of height s , by Krull’s principal ideal theorem. And it is generated by polynomials of degree $\ell' \leq \max\{\ell, (\ell - 1)^s\}$, since g has degree at most $(\ell - 1)^s$. So we may apply, by induction on the height, the assertion of the theorem to I' : There exists an exact formal solution $\widehat{y}(x)$ of $f(x, y) = g(x, y) = 0$ with $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^c$. This suggests defining $\beta_{n,m,\ell}(c)$ in this case as $\beta'_{n,m,\ell'}(c)$ with $\ell' = \max\{\ell, (\ell - 1)^{\lambda(n,m,\ell)}\}$.

(ii) We now assume that $\overline{y}(x)$ is not an approximate solution of $g(x, y) = 0$ up to degree $\beta'(c)$. Denote by $f^* = (f_1, \dots, f_s)$ the vector whose components generate the ideal I^* .

The intersection $X \cap Y$ of the component X of X^* with the union $Y = V(J)$ of the other components of X^* is contained in the singular locus $\text{Sing}(X^*)$ of X^* . As X^* is pure dimensional of codimension s , the singular locus is defined by the ideal of $(s \times s)$ -minors of f^* . It follows that the minor g vanishes identically on $X \cap Y$. As our ground field \mathbb{K} is assumed to be algebraically closed, Hilbert’s Nullstellensatz tells us that g belongs to the radical of the ideal $I + J$ defining $X \cap Y$. By Hermann’s result, the degree of $I + J$ is bounded by a function λ in ℓ, m , and n . Moreover, there exists an integer u bounded by λ so that $g^u \in I + J$.

Choose $\beta(c) = \beta'(c')$ with $c' = u \cdot \beta'(c) + 1$, where u is such that $g^u \in I + J$ and $\beta'(c) = \beta'_{n,m,\ell''}(c)$ with $\ell'' = \lambda(n, m, \ell)$. As $\overline{y}(x)$ is not an approximate solution of $g(x, y) = 0$ up to degree $\beta'(c)$, it is not an approximate solution of $g^u(x, y) = 0$ up to degree $u \cdot \beta'(c)$. But $g^u \in I + J$ and $\overline{y}(x)$ is an approximate solution of $f(x, y) = 0$ (say, for I) up to degree $\beta(c) \geq u \cdot \beta'(c) + 1$. Therefore there exists an element $h \in J$ so that $\overline{y}(x)$ is not an approximate solution of $h(x, y) = 0$ up to degree $u \cdot \beta'(c)$.

Recall that $\overline{y}(x)$ is an approximate solution of $f(x, y) = 0$ up to degree $\beta(c) \geq \beta'(c)$. Applying the theorem to I^* with $\beta'(c')$, we get an exact solution $\widehat{y}(x)$ for I^* with $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^{c'}$ with $c' = u \cdot \beta'(c) + 1$. Now observe that $h \cdot I \subset J \cdot I \subset I \cap J = I^*$. Hence $\widehat{y}(x)$ is also an exact solution for $h \cdot f$. But $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^{c'}$ and $h(x, \overline{y}(x)) \not\equiv 0$ modulo $(x)^{u \cdot \beta'(c)}$, so that $h(x, \widehat{y}(x)) \neq 0$. It follows that $f(x, \widehat{y}(x)) = 0$ as required. From $c' \geq c$ follows also that $\widehat{y}(x) \equiv \overline{y}(x)$ modulo $(x)^c$.

Combining (i) and (ii), we see that $\beta_{n,m,\ell}(c)$ has to be chosen larger than or equal to $\beta'_{n,m,\ell'}(c)$, with $\ell' = \max\{\ell, (\ell - 1)^{\lambda(n,m,\ell)}\}$, and also $\beta'_{n,m,\ell''}(c')$, with $c' = u \cdot \beta'_{n,m,\ell''}(c) + 1$ and $\ell'' = \lambda(n, m, \ell)$.

Step 3. Proof in the complete intersection case. We can now assume we have an ideal I of height r which is generated by r polynomials f_1, \dots, f_r for which an $(r \times r)$ -minor g of $\partial_{xy}f$ does not belong to I . By the same reasoning as in case (i) of Step 2, we may suppose additionally that the given approximated solution $\overline{y}(x)$ of $f(x, y) = 0$ is so that $g(x, \overline{y}(x))$ is nonzero up to degree $\beta(c) + 1$.

We derive $f(x, \overline{y}(x)) \equiv 0$ modulo $(x)^{\beta(c)}$ with respect to the x -variables. We get the congruence

$$\partial_{x_i} f(x, \overline{y}(x)) \equiv - \sum_{k=1}^r \partial_{x_i} \overline{y}_k(x) \cdot \partial_{y_k} f(x, \overline{y}(x))$$

modulo $(x)^{\beta(c)-1}$. This shows that we can choose the minor g from the relative Jacobian matrix $\partial_y f$.

We then follow the proof of the convergent and algebraic cases until we arrive at the equivalent system of polynomial equations in $x' = (x_1, \dots, x_{n-1})$ in new unknown series $w_0(x'), \dots, w_{d-1}(x')$. We have already mentioned that the polynomial vector h of Step 3 of the proof will have a degree bounded by a function in the degree of f and m , namely by $\deg(f) + m \cdot (\deg(f) - 1)^{m-1} \leq \ell + m \cdot (\ell - 1)^{m-1}$.

Now let $\beta'_{n-1,m,\ell}(c)$ denote the function of the theorem given by induction in dimension $n - 1$. It then suffices to set $\beta_{n,m,\ell}(c) = \beta'_{n-1,m,\ell'}(c)$ with $\ell' = \ell + m \cdot (\ell - 1)^{m-1}$ to establish the case of dimension n . This concludes the proof of the uniform strong approximation theorem. \square

(e) The proof of the strong approximation theorem for textile maps only relies on the Noetherianity of the polynomial ring and on Chevalley's theorem on the images of constructible sets under polynomial maps [Ma, p. 42]. It is this latter result which requires that we work over algebraically closed fields.

We adapt an argument that was used by Mori to prove a theorem of Gurjar and Parameswaran on the boundedness of Milnor numbers in families of hyperplane sections [Mo], [Gu]. Set $M = \mathbb{K}[[x]]^m$ and $M_k = M/(x)^k M$ for $k \geq 0$, with projection maps $\pi_k : M \rightarrow M_k$. We identify the elements of M_k with vectors of polynomials of degree $< k$ and consider M_k as affine space $\mathbb{A}_{\mathbb{K}}^{mnk}$. As G is continuous with respect to the Krull topology, there exists an increasing sequence of natural numbers c_k tending to infinity such that if y and y' are congruent modulo $(x)^k$, then $G(y)$ and $G(y')$ are congruent modulo $(x)^{c_k}$. Fix such a sequence. We consider the subsets

$$V_k = \pi_k(\{y \in M, \text{ord } G(y) \geq c_k\}).$$

of M_k . They are nonempty by assumption. As G is textile, these are constructible subsets of M_k .³ This will be used later on. For $\ell \geq k$, let $\pi_{\ell,k} : M_\ell \rightarrow M_k$ denote the canonical projections. If $y \in M_\ell$ has image $\bar{y} \in M_k$, then y and \bar{y} are congruent modulo $(x)^k$, and therefore $G(y)$ and $G(\bar{y})$ are congruent modulo $(x)^{c_k}$. In particular, if y belongs to V_ℓ and hence $G(y)$ has order $\geq c_\ell$, then, due to $c_\ell \geq c_k$, we get that $G(\bar{y})$ has order $\geq c_k$. Thus $\bar{y} \in V_k$. This shows that $\pi_{\ell,k}$ maps V_ℓ into V_k , for $\ell \geq k$.

For each k , the descending chain of Zariski-closures $\overline{\pi_{\ell,k}(V_\ell)}$ in V_k , with $\ell \geq k$, stabilizes by the Noetherianity of the Zariski-topology for large ℓ , say for $\ell \geq \lambda_k$. For $k \geq 0$ and $\ell \geq \lambda_k$, we set

$$W_k = \overline{\pi_{\ell,k}(V_\ell)} \subset V_k.$$

Let $\ell \geq k$, and write $W_\ell = \overline{\pi_{\ell',\ell}(V_{\ell'})}$ for some $\ell' \geq \ell$. Choosing $\ell' \geq \lambda_k$, we get

$$\overline{\pi_{\ell,k}(W_\ell)} = \overline{\pi_{\ell,k}(\overline{\pi_{\ell',\ell}(V_{\ell'})})} = \overline{\pi_{\ell,k}(\pi_{\ell',\ell}(V_{\ell'}))} = \overline{\pi_{\ell',k}(V_{\ell'})} = W_k.$$

Hence the induced projections $\tau_{\ell,k} : W_\ell \rightarrow W_k$ are dominant for $\ell \geq k$ (i.e., have dense images).

Let us now put ourselves in the situation of the theorem. We claim that taking $k = 0$, any value $\ell \geq c_{\lambda_0}$ will be sufficient for the conclusion of the theorem. By assumption, there then exists an element $\bar{y}(x) \in M$ such that $G(\bar{y}(x)) \equiv 0$ modulo $(x)^\ell$, say $\text{ord } G(\bar{y}(x)) \geq \ell \geq c_{\lambda_0}$. This implies that the image $\pi_{\lambda_0}(\bar{y}(x))$ of $\bar{y}(x)$ in M_{λ_0} belongs to V_{λ_0} . By the definition of λ_0 , the image $\pi_{\lambda_0,0}(\pi_{\lambda_0}(\bar{y}(x))) = \pi_0(\bar{y}(x))$

³A subset is *constructible* if it is a finite union of differences of Zariski-closed sets. We invoke here Chevalley's theorem.

of $\bar{y}(x)$ in M_0 belongs to W_0 . It follows that W_0 and, consequently, all W_k are nonempty.

We now construct the exact solution $y(x)$ by lifting a sufficiently general element of W_0 suitably to all W_ℓ and then to M . The element $\bar{y}(x)$ is not used for this lifting.

We have seen above that the images $\tau_{\ell,0}(W_\ell)$ are Zariski-dense in W_0 . As our ground field is algebraically closed, these images contain a Zariski-open subset of W_0 , by Chevalley's theorem. Therefore the countable intersection $\bigcap_{\ell \geq 0} \tau_{\ell,0}(W_\ell)$ is again Zariski-dense in W_0 .

As the image of $V_{\ell'}$ in M_k is constructible and Zariski-dense in W_k for $\ell' \geq \lambda_k$, the same argument shows that the intersection of all images of the sets $V_{\ell'}$ in W_k , with $\ell' \geq \lambda_k$, is dense in W_k for all k . Call this intersection U_k , say $U_k = \bigcap_{\ell' \geq \lambda_k} \pi_{\ell',k}(V_{\ell'})$. Let λ_ℓ be the corresponding index for W_ℓ . We may assume that $\lambda_\ell \geq \lambda_k$. As the sequence of images $\pi_{\ell',k}(V_{\ell'})$ is decreasing in W_k , we also have $U_k = \bigcap_{\ell' \geq \lambda_\ell} \pi_{\ell',k}(V_{\ell'})$. This shows that the restriction $U_\ell \rightarrow U_k$ of $\pi_{k,\ell}$ is onto for $\ell \geq k$. Therefore any element of W_k which lies in the dense subset U_k lifts successively to all $U_\ell \subset V_\ell$ (successively in the sense that the limit exists as a formal power series). This limit will give, by definition of V_ℓ and since c_ℓ tends to infinity, an exact solution $y(x)$ of $G(y(x)) = 0$. The proof of the strong approximation theorem for textile maps is completed. \square

5. WEIERSTRASS DIVISION THEOREMS

In the division theorems below, x and z denote two sets of variables $x = (x_1, \dots, x_n)$ and $z = (z_1, \dots, z_m)$, and x' is set equal to (x_1, \dots, x_{n-1}) . Fix $d \in \mathbb{N}$. Let $w_i = (w_{i1}, \dots, w_{im})$ denote additional sets of variables, for $i = 1, \dots, d$, and write w for the collection of all w_i . Denote by

$$z(w, x_n) = \sum_{i=1}^d w_i \cdot x_n^i \in \mathbb{K}[w][x_n]^m,$$

with nonitalic letter z , the universal polynomial vector in x_n of degree $\leq d$ with variable coefficients $w_i = (w_{i1}, \dots, w_{im})$ and vanishing for $x_n = 0$. Any choice of formal power series vectors $w_i(x') \in \mathbb{K}[[x']]^m$ then induces, writing $w(x')$ for the collection of all $w_i(x')$, a polynomial vector

$$z(x) = z(w(x'), x_n) = \sum_{i=1}^d w_i(x') \cdot x_n^i \in \mathbb{K}[[x']][x_n]^m$$

in x_n of degree $\leq d$ with coefficients the formal power series vectors $w_i(x')$.

Classical Weierstrass Division Theorem. *Let g be a nonzero formal, convergent, or algebraic power series in n variables x_1, \dots, x_n . Assume that g is x_n -regular of order e . Then there exist, for any formal, convergent, or algebraic power series f , unique formal, convergent, or algebraic power series p and q , with q polynomial in x_n of degree $< e$, such that*

$$f(x) = p(x) \cdot g(x) + q(x).$$

The algebraic case was originally proven by Lafon [Laf1], [Laf2]; see also [Hir], [AMR], [ACH], [Rui, Prop. 5.6]. We give below an elementary, more constructive proof of the theorem in all three settings.

Proof. (a) We first treat the formal case. Write $g = x_n^e + h$ with $h \in [(x') + (x_n^{e+1})] \cdot \mathbb{K}[[x]]$. Consider the map

$$w : \mathbb{K}[[x]] \times \mathbb{K}[[x']][x_n]_{<e} \rightarrow \mathbb{K}[[x]], (p, q) \rightarrow p \cdot g + q.$$

We have to show that w is an isomorphism. For this it suffices to decompose w into $w = u - v$ where $u(p, q) = p \cdot x_n^e + q$ is trivially an isomorphism, and $v(p, q) = -p \cdot h$. Hence it suffices to show that $wu^{-1} = \text{Id}_{\mathbb{K}[[x]]} - vu^{-1}$ is an isomorphism. Equip the variables x_1, \dots, x_{n-1} with weight 1 and x_n with weight $\frac{1}{e+1}$. Then x_n^e is the monomial with smallest weight of the expansion of g . It follows that vu^{-1} increases the weighted order of power series, which, in turn, shows that the geometric series $\sum_{k=0}^{\infty} (vu^{-1})^k$ induces a well-defined map $\mathbb{K}[[x]] \rightarrow \mathbb{K}[[x]]$. It is clearly the inverse to wu^{-1} .

(b) In the convergent case, write $\mathbb{K}\{x\}$ as the union of the rings $\mathbb{K}\{x\}_r$ of power series whose radius of convergence is at least $r > 0$. The rings $\mathbb{K}\{x\}_r$ are Banach spaces with the norm $|\sum_{\alpha} c_{\alpha} x^{\alpha}| = \sum_{\alpha} |c_{\alpha}| \cdot r^{|\alpha|}$. It is easy to show that the restriction of vu^{-1} to $\mathbb{K}\{x\}_r$ maps $\mathbb{K}\{x\}_r$ into itself for sufficiently small $r > 0$ and has operator norm < 1 . Therefore $\sum_{k=0}^{\infty} (vu^{-1})^k$ induces a well-defined map $\mathbb{K}\{x\}_r \rightarrow \mathbb{K}\{x\}_r$ which will be the inverse to the restriction of wu^{-1} to $\mathbb{K}\{x\}_r$.

(c) The algebraic case is more complicated. We give the main line of arguments, for detail see [AMR], [ACH]. It can be shown, by the Artin–Mazur lemma [AM, p. 88], that for given algebraic series f and g there exists a polynomial vector $H = (H_1, \dots, H_p)$ in variables x_1, \dots, x_n and y_1, \dots, y_p for which $H(0, 0) = 0$ and $\partial_y H(0, 0)$ is an invertible matrix, and such that the unique power series solution $h(x) = (h_1(x), \dots, h_p(x))$ to $H(x, y(x)) = 0$ given by the implicit function theorem is algebraic and so that f and g can be expressed as polynomials $f(x) = F(x, h(x))$ and $g(x) = G(x, h(x))$ in $h(x)$ [AM], [AMR], [ACH]. These polynomials H , F , and G are called the *codes* of f , respectively g . Set

$$B_i = y_i + \sum_{j=0}^{e-1} u_{ij} \cdot x_n^j$$

and

$$B = x_n^e + \sum_{j=0}^{e-1} u_j \cdot x_n^j$$

with new variables u_{ij} and u_j , for $i = 1, \dots, m$. We may consider these polynomials equipped with the leading terms y_i and x_n^e . It is not hard to see that the ideals in the ring of algebraic series generated by $y_1 - h_1, \dots, y_p - h_p$ and g , respectively by H_1, \dots, H_p and G , coincide. The division of the polynomials H_k and G by B_1, \dots, B_m and B yields polynomial remainders

$$R_k(u_{ij}, u_j, x) = \sum_{\ell=0}^{e-1} U_{k\ell}(u_{ij}, u_j, x') \cdot x_n^{\ell},$$

$$R(u_{ij}, u_j, x) = \sum_{\ell=0}^{e-1} U_{\ell}(u_{ij}, u_j, x') \cdot x_n^{\ell},$$

which do not depend on y . One then shows that the system given by $U_{k\ell}(u_{ij}, u_j, x') = 0$ and $U_{\ell}(u_{ij}, u_j, x) = 0$, for $k = 1, \dots, m$ and $\ell = 0, \dots, e - 1$, satisfies with respect to u_{ij} and u_j the assumptions of the implicit function theorem. Therefore there

exist unique algebraic series $u_{ij}(x')$ and $u_j(x')$ such that $U_{k\ell}(u_{ij}(x'), u_j(x'), x') = 0$ and $U_\ell(u_{ij}(x'), u_j(x'), x) = 0$ for all k and ℓ . By construction, the series $b(x) = B(u_j(x'), x_n)$ is then algebraic and generates in $\mathbb{K}[[x]]$ the same ideal as $g(x)$, for being its Weierstrass normal form.

It now suffices to divide F polynomially by B_1, \dots, B_m and B ,

$$F = \sum_{i=1}^m A_i \cdot B_i + A \cdot B + C$$

with polynomials A_1, \dots, A_m, A and C in $\mathbb{K}[u_{ij}, u_j, x, y]$, where $C = C(u_{ij}, u_j, x)$ does not depend on y and is polynomial in x_n of degree $\leq e$. Upon replacing in this equation y_i by $h_i(x)$, respectively u_{ij} and u_i by $u_{ij}(x')$ and $u_j(x')$, one gets

$$f = \sum_{i=1}^m a_i \cdot b_i + a \cdot b + c$$

with algebraic series a_1, \dots, a_m, a and c obtained from A_1, \dots, A_m, A and C by substitution. Here, c is a polynomial in x_n of degree $\leq e - 1$. To conclude the proof, it suffices to observe that b_1, \dots, b_m are algebraic multiples of b , hence of g , so that the required Weierstrass division $f = p \cdot g + q$ follows by substitution. Note that this proof is constructive as it gives the codes of the quotient p and the remainder q from the codes of f and g . □

Universal Weierstrass Division Theorem. *Fix d , and let $z(w, x_n)$ denote the universal polynomial in x_n of degree d as before. Assume given a formal power series $g(x, z)$ so that $g(x, z(w, x_n))$ is nonzero and x_n -regular of order e .*

(a) *For every formal power series $f(x, z)$ there exist a unique formal power series $p \in \mathbb{K}\{x, w\}$ and a unique polynomial $q \in \mathbb{K}\{x', w\}[x_n]$ in x_n of degree $< e$ with formal power series coefficients in x' and w such that*

$$f(x, z(w, x_n)) = p(x, w) \cdot g(x, z(w, x_n)) + q(x, w)$$

is the Weierstrass division of $f(x, z(w, x_n))$ by $g(x, z(w, x_n))$.

(b) *For every evaluation $z(x) = z(w(x'), x_n) \in \mathbb{K}[[x']][x_n]^m$ of $z(w, x_n)$ at formal power series vectors $w_i(x')$ for which $g(x, z(x))$ is nonzero and x_n -regular of order e , the induced decomposition*

$$f(x, z(x)) = p(x, w(x')) \cdot g(x, z(x)) + q(x, w(x'))$$

is the formal Weierstrass division of $f(x, z(x))$ by $g(x, z(x))$.

(c) *The same statements as in (a) and (b) hold for convergent, respectively algebraic, power series f and g , in which case q and p are again convergent, respectively algebraic.*

(d) *If f and g are polynomials of total degree in x and y less than or equal to ℓ , the quotient p and the remainder q are again polynomials and the total degree in x' and w of the coefficients of q is bounded by $\ell + \ell^2$.*

Proof. Assertion (a) follows from the classical formal Weierstrass division theorem in the case of two sets of variables x and w . Assertion (b) follows from (a) since $w(x')$ does not depend on x_n , so that $q(x, w(x'))$ is again a polynomial in x_n of degree $< e$. The convergent and algebraic versions of the classical Weierstrass division theorem then yield assertion (c).

Finally, for assertion (d), observe that the degree of the polynomial $f(x, z(w, x_n))$ in x_n is bounded by $t = \deg_x f + d \cdot \deg_y f$. The Weierstrass division of $f(x, z(w, x_n))$

by $g(x, z(w, x_n))$ coincides with the respective polynomial division with respect to the leading monomial x_n^e of $g(x, z(w, x_n))$. It consists in replacing e th powers of x_n by lower degree powers. Such substitutions occur at most $t - e$ times. In each step the degree of the intermediate remainders in the w -variables increases by 1, so that the final remainder q will have degree in w at most $t - e$. Similarly, the degree of q in x' will be less than or equal to $\deg_{x'} f + (t - e) \cdot \deg_{x'} g$. Combining these estimations yields for the total degree of q in x' and w the upper bound

$$t - e + \deg_{x'} f + (t - e) \cdot \deg_{x'} g = \deg_{x'} f + (t - e) \cdot (\deg_{x'} g - 1).$$

This can be bounded from above by $\ell + \ell^2$, giving the required estimate. □

6. ALGEBRAIC POWER SERIES

These series play a prominent role in the approximation theorems, especially in the nested case. They also appear in Azumaya’s and Nagata’s theory of Henselian rings and Grothendieck’s étale topology. Above all, they represent a fascinating research topic with many facets, linking commutative algebra with complex and asymptotic analysis, diophantine approximation, counting problems in combinatorics, and specific phenomena over fields of characteristic p . We propose giving a brief (and rather incomplete) account of them. Our main sources are lectures by Bostan [Bos], the survey paper [BD] of Banderier and Drmota, and an article by Adamczewski and Bell [AB].

A formal power series $y(x)$ in n variables x_1, \dots, x_n over \mathbb{K} is called *algebraic* if there exists a (nonzero) polynomial $P(x, t) = c_d(x)t^d + \dots + c_1(x)t + c_0(x) \in \mathbb{K}[x, t]$ so that $P(x, y(x)) = 0$. Typical examples are *rational power series* as, e.g., fractions $\frac{1}{1+z(x)}$, or roots $\sqrt{1+z(x)}$ with $z(x)$ a polynomial with constant term $z(0) \neq -1$. Functions such as $\frac{1}{x}$ or \sqrt{x} are not algebraic (since they are not even formal power series). The components of the local formal inverse at 0 of a polynomial map $F : \mathbb{K}^n \rightarrow \mathbb{K}^n$ with $F(0) = 0$ and $\det(\partial_x F(0)) \neq 0$ are algebraic series [LT], as are the components of the solutions $y(x)$ vanishing at 0 of an implicit polynomial system $G(x, y) = 0$ with $G(0, 0) = 0$ and $\partial_y G(0, 0)$ invertible, for a polynomial map $G : \mathbb{K}^{n+m} \rightarrow \mathbb{K}^m$. The set of algebraic series forms a subring $\mathbb{K}\langle x \rangle$ of $\mathbb{K}[[x]]$ which is closed under derivation but not under composition with other algebraic series.

The logarithmic series $\log(1 - x) = \sum_{i \geq 1} \frac{x^i}{i}$ and the exponential series $\exp(x) = \sum_{i \geq 1} \frac{x^i}{i!}$ in one variable ($\text{char } \mathbb{K} = 0$) are clearly *transcendental* (i.e., not algebraic) as well as the *lacunary series* $h(x) = \sum_{i=0}^{\infty} x^{2^i}$. The latter is, however, algebraic over a field \mathbb{K} of characteristic 2 (apply the linearity of squaring to the *Mahler equation* $h(x^2) - h(x) = x$ satisfied by h).

In general, it is easier to prove that a series is transcendental than to show that it is algebraic. Over a valued field, an algebraic series is automatically convergent. So the coefficients cannot grow faster than polynomially. Otherwise the series is transcendental. Similarly, over a valued field of characteristic 0, a series $y(x)$ is also transcendental if the coefficients tend too fast to 0 (as is the case with $\exp(x)$).

We list a few results (mostly for series in one variable with coefficients in \mathbb{Q}, \mathbb{C} , or a field of positive characteristic).

Eisenstein 1852 [Eis], Heine 1853 [Hei]: *Univariate algebraic series $\sum c_i t^i$ over \mathbb{Q} admit a common “denominator”—there is a $d \in \mathbb{N}$ so that $c_i d^i \in \mathbb{Z}$ for all $i \geq 1$.*

In particular, the denominators of the coefficients c_i have only finitely many prime divisors. Eisenstein was the first to observe the phenomenon. He proved the statement in the case where the minimal polynomial of the series satisfies at 0 the assumption of the implicit function theorem; Heine then did the general case.

Abel 1827 [Abe, p. 287], Cockle 1860 [Coc], Harley 1862 [Harl], Tannery 1874 [Tan], [Sta], [Lip]: *Univariate algebraic series $h(t) = \sum c_i t^i$ are D -finite, i.e., their coefficients satisfy a linear recursion $c_i = a_1 c_{i-1} + \dots + a_d c_{i-d}$, for $i \geq d$, with a_i rational functions in i . Equivalently, $h(t)$ satisfies a linear differential equation with polynomial coefficients.*

If $f(x) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha x^\alpha$ is a formal power series in variables x_1, \dots, x_n , define its *diagonal* as the series in one variable z given by $\text{diag}(f)(z) = \sum_i c_{(i, \dots, i)} z^i$ [Sta], [Lip].

Pólya 1922 [Pol]: *Diagonals of bivariate rational power series over \mathbb{C} are algebraic.*

This can be seen as follows. Let $f(x, y)$ be rational. Then we can interpret its diagonal $\text{diag}(f)(z)$ as the coefficient of the series $\frac{1}{x} f(x, \frac{z}{x})$ at x^{-1} . This coefficient is an integral, namely

$$\text{diag}(f)(z) = \frac{1}{2\pi i} \int_{|x|=\varepsilon} \frac{1}{x} f\left(x, \frac{z}{x}\right) dx.$$

The integral can be computed via residues and is hence algebraic. The striking fact is that Pólya's statement is actually an equivalence in characteristic p .

Furstenberg 1967 [Fur, Thm. 1 and 2]: *Over a field of positive characteristic $p > 0$, the diagonal of a rational function in several variables is an algebraic series, and every algebraic power series in one variable over \mathbb{C} is the diagonal of a bivariate rational function.*

Example ([Bos, part II]). Let $f(z)$ be algebraic in one variable over \mathbb{Q} , with minimal polynomial $P(z, y)$. Assume that $P(0, 0) = 0$ and $\partial_y P(0, 0) \neq 0$, i.e., that P satisfies at 0 the assumption of the implicit function theorem. Then

$$f(z) = \text{diag} \left(y^2 \cdot \frac{\partial_y P(xy, y)}{P(xy, y)} \right).$$

The equivalence need not hold for more than two variables, as is seen from the diagonal of $\frac{1}{1-x-y-z}$, which equals the transcendental hypergeometric series ${}_2F_1(1/3, 2/3; 1; 27z)$.

Deligne extended Pólya's and Furstenberg's results to diagonals of algebraic series.

Deligne 1983 [Del], [DL3], [Hara], [ShW, Chr]: *Over a field of positive characteristic, the diagonal of an algebraic series in n variables is again algebraic.*

Denef and Lipshitz 1987 [DL3]: *Regardless of the characteristic of the ground field, every univariate algebraic function is the diagonal of a bivariate rational function.* (This is even true for any number of variables, taking the large diagonal.)

Example ([AB, p. 3]). In characteristic 0, the statement of Deligne's result is no longer valid. The real series

$$h(x) = \sum_{i=0}^{\infty} \frac{1}{2^{4i}} \binom{2i}{i}^2 x^i = \frac{2}{\pi} \int_0^{\pi/2} \frac{dt}{\sqrt{1-x \sin^2 t}}$$

can be realized as the diagonal of two rational power series in four, respectively three, variables $\frac{2}{2-u-v} \cdot \frac{2}{2-wz}$ and $\frac{4}{4-(u+v)(1+w)}$, but is not algebraic.

We now come briefly to Grothendieck's famous and wide open p -curvature conjecture (unpublished; see [Kat1], [Kat2], [dVRSZ]). It is on the reduction of systems of ordinary differential equations modulo p .

Grothendieck 1969: *Let $A \in \mathbb{Q}(t)^{m \times m}$ be a square matrix whose entries $a_{ij}(t)$ are univariate rational functions over \mathbb{Q} . Consider the system of ordinary linear differential equations*

$$(*) \quad y'(t) = A(t) \cdot y(t),$$

for unknown functions $y(t) = (y_1(t), \dots, y_m(t))$. Then $(*)$ has a fundamental system of solutions in the algebraic closure of $\mathbb{Q}(x)$ if and only if, for almost all primes p , the reduction of $(*)$ modulo p has a fundamental system of solutions consisting of algebraic power series over \mathbb{F}_p (or, equivalently, rational power series over \mathbb{F}_p).

Here, by a fundamental system we understand a basis of the solution space. One may take, instead of $(*)$, also a single ordinary differential equation in $y(t) \in \mathbb{Q}[[t]]$ with polynomial coefficients,

$$(\#) \quad a_k(t) \cdot y^{(k)} + a_{k-1}(t) \cdot y^{(k-1)} + \dots + a_1(t) \cdot y' + a_0(t) \cdot y = 0.$$

The name p -curvature stems from the following equivalent statement of the conjecture. Define recursively matrices A_k by $A_{k+1} = A'_k + A_k \cdot A$, starting with the identity matrix $A_0 = \mathbb{1}_m$ and letting $A' = \partial_t A$ be the derived matrix. For p prime, the matrix A_p is called the p -curvature of the system $(*)$. Then the assertion of the conjecture is equivalent to the following: *the matrices A_p are congruent to 0 modulo p for almost all primes p* . Said differently, the p -curvature is the p th iterate $\Psi_p = (\frac{d}{dt} + A)^p$, taken modulo p . This is a linear operator, and $\Psi_p = A_p$. The conjecture is open. The case of Picard–Fuchs equations was solved by Katz in 1982 [Kat1], [Kat2]. There is also a discrete version of the conjecture, replacing differentiation by the difference operator, by di Vizio [dV]. The long proof relies on André's theory of G -functions.

Within commutative algebra, the ring of algebraic power series can be alternatively defined as the *Henselization* of the localization $\mathbb{K}[x]_{(x)}$ at 0 of the polynomial ring at the maximal ideal $(x) = (x_1, \dots, x_n)$. This goes as follows.

A Noetherian local ring (A, m) is called *Henselian* if every univariate polynomial $P(t)$ with coefficients in A admitting a simple approximate root $\bar{\alpha}$ in A modulo the maximal ideal m of A , say $P(\bar{\alpha}) \equiv 0$ modulo m , also admits a (simple) root α in A lifting the approximate root, i.e., satisfying $f(\alpha) = 0$ and $\alpha \equiv \bar{\alpha}$ modulo m . It is equivalent to saying that every monic polynomial $P(t)$ which factors modulo m into two coprime monic polynomials \bar{Q} and \bar{R} , say $P \equiv \bar{Q} \cdot \bar{R}$ modulo m , also factors exactly, $P = R \cdot Q$, with Q and R monic polynomials satisfying $R \equiv \bar{R}$ and $Q \equiv \bar{Q}$ modulo m .

A local ring (A, m) always admits a Henselization, i.e., a smallest ring extension $A \subset A^h$ for which A^h is Henselian [Na]. More precisely, A^h is given by the following universal property: every local ring homomorphism $A \rightarrow B$ to a Henselian ring B extends uniquely to $A^h \rightarrow B$. As an example take for instance for A the localization $\mathbb{K}[x]_{(x)}$ of the polynomial ring $\mathbb{K}[x]$ at the maximal ideal (x) , one obtains for A^h precisely the ring $\mathbb{K}\langle x \rangle$ of algebraic power series.

In a Henselian ring A , the characterization of algebraic series by implicit polynomial equations $G(y) = 0$ in several variables with $G \in A[y]^m$ and $\partial_y G(0)$ invertible is also known as the *multivariate Hensel lemma*; cf. [EGA, 18.5.11], [Ra2]. The characterization uses the concept of the *normalization* of an algebraic variety. It can also be expressed in the language of étale coverings.

Theorem (Artin and Mazur, [AM, p. 88]). *A formal power series $h(x_1, \dots, x_n)$ with $h(0) = 0$ is algebraic if and only if it is the first (or any other) component of the unique solution $y(x)$ with $y(0) = 0$ of an implicit equation $F(x, y) = 0$ for a polynomial map $F : \mathbb{K}^{n+m} \rightarrow \mathbb{K}^m$ with $F(0, 0) = 0$ and $\det(\partial_y F(0, 0)) \neq 0$.*

Proof. One direction is easy: Extend F to $G = (x, F) : \mathbb{K}^{n+m} \rightarrow \mathbb{K}^{n+m}$. It satisfies at 0 the assumptions of the inverse function theorem, $G(0, 0) = 0$, and $\det(\partial_{xy} G(0, 0)) \neq 0$. By the inverse function theorem its inverse $H(x, y) = G^{-1}(x, y)$ at 0 has algebraic components [LT]. We may write $H(x, y) = (x, \tilde{H}(x, y))$. Setting $y = 0$ provides the solution $y(x) = \tilde{H}(x, 0)$ to $F(x, y) = 0$.

Conversely, let h be given with minimal polynomial $P(x, y)$. We notice for later use that P is irreducible. Denote by X the algebraic hypersurface in $\mathbb{A}_{\mathbb{K}}^{n+1}$ defined by $P = 0$. As $P(x, h(x)) = 0$, we see that X contains the graph of h (think either of $\mathbb{K} = \mathbb{C}$ and that h is holomorphic in a euclidean neighborhood of \mathbb{C}^n or define the graph scheme-theoretically). We may associate to X the germ $X_0 = (X^{an}, 0)$ at 0 of the analytic variety X^{an} defined by P (if \mathbb{K} is a valued field) or, for arbitrary fields \mathbb{K} , the formal neighborhood X_0 of X at 0 (defined by the ring $\mathbb{K}[[x, y]]/(P)$). In both cases, the graph of h lies inside one analytic component C of X_0 , which, by comparison of dimensions, must hence be smooth. Nevertheless, X may be singular at 0, since several analytic components could meet.

The idea now is to separate these components by *normalization*. It is given by taking the integral closure A of the coordinate ring $\mathbb{K}[x, y]/(P)$ inside its quotient field (which is well-defined since P is irreducible). This closure is again a finitely generated \mathbb{K} -algebra [ZS], of Krull dimension n as X . It can therefore be written as $A = \mathbb{K}[z_1, \dots, z_p]/I$, for some ideal I . This ideal defines a variety Y in $\mathbb{A}_{\mathbb{K}}^p$, called the normalization of X . The inclusion $\mathbb{K}[x, y]/(P) \subset A$ defines a (finite) morphism $\pi : Y \rightarrow X$. By the universal property of normalization, the map $\gamma : x \rightarrow (x, h(x))$ from $\mathbb{A}_{\mathbb{K}}^n$ to X lifts to a map $\tilde{\gamma} : \mathbb{A}_{\mathbb{K}}^n \rightarrow Y$, i.e., so that $\pi \circ \tilde{\gamma} = \gamma$. By a precursor of Zariski's main theorem [Za], one knows that Y is analytically irreducible at every point above 0. In particular, the map π separates the analytic components of X at 0. So there is a unique point $0'$ in $C' = \pi^{-1}(C)$ lying over 0. Without loss of generality, we may assume that $0'$ is the origin of $\mathbb{A}_{\mathbb{K}}^p$, and that $\pi : Y \rightarrow X$ is given by the restriction to Y of the projection map $\mathbb{A}_{\mathbb{K}}^p \rightarrow \mathbb{A}_{\mathbb{K}}^{n+1}$ on the first $n+1$ components. In particular, $\tilde{\gamma}$ has first $n+1$ components $(x_1, \dots, x_n, h(x))$. As C is smooth at 0, this now also holds for C' at $0'$.

To summarize, we have shown that the ideal I defines an algebraic subvariety Y of $\mathbb{A}_{\mathbb{K}}^p$ of codimension $p-n$ which is smooth at $0'$. The Jacobian criterion of smoothness now tells us that there are polynomial generators Q_1, \dots, Q_{p-n} of I so that the matrix $(\partial_{z_i} Q_j)_{i,j=1, \dots, p-n}$ has rank $p-n$ at 0. In this situation, the implicit function theorem applies. It shows that there is a (unique) formal power series vector (actually, with components algebraic series) $(h_1, \dots, h_{p-n}) \in \mathbb{K}[[z_{p-n+1}, \dots, z_p]]^{p-n}$ so that, setting $z' = (z_{p-n+1}, \dots, z_p)$, one has $h_i(0) = 0$ and $Q_j(h_1(z'), \dots, h_{p-n}(z'), z') = 0$ for all i and j .

It now suffices to define $F : \mathbb{A}_{\mathbb{K}}^p \rightarrow \mathbb{A}_{\mathbb{K}}^p$ as $F(z) = (z_1, \dots, z_n, Q_1(z'), \dots, Q_{p-n}(z'))$ to recognize h_1 and hence h as a component of the inverse of F . This proves the theorem. \square

7. FORMAL AND ANALYTIC RELATIONS BETWEEN CONVERGENT SERIES

Let $f_1(x), \dots, f_m(x)$ be given convergent power series vanishing at 0 in variables $x = (x_1, \dots, x_n)$ over a valued field \mathbb{K} . We wish to compare the formal and analytic relations $r = r(y_1, \dots, y_m)$ between the f_i , i.e., series $r(y_1, \dots, y_m)$ so that

$$r(f_1(x), \dots, f_m(x)) = 0.$$

Grothendieck asked whether the ideal of formal relations $I = \{r, r(f) = 0\}$ is generated by the analytic ones. Gabrielov showed by means of a tricky example that this is not the case [Ga1], [Ga2], [Izu]. We will have a closer look at Gabrielov’s construction and explain the clue behind it. The details can be found in [ACHK].

The equation above is equivalent to saying that $r(y)$ belongs to the ideal generated by the series $y_i - f_i(x)$ for $i = 1, \dots, m$. Therefore, there exist power series $a_1(x, y), \dots, a_m(x, y)$ such that

$$r(y) = \sum_{i=1}^m a_i(x, y) \cdot (y_i - f_i(x)).$$

Here, the series a_i are allowed to depend on both x and y , whereas the series r must be independent of x . One may then ask more generally, Given analytic functions e and f_1, \dots, f_m in n variables x_1, \dots, x_n such that the linear presentation

$$e(x) = \sum_{i=1}^m \widehat{a}_i(x) \cdot f_i(x)$$

holds with formal power series $\widehat{a}_i(x)$ depending only on the variables x_1, \dots, x_{n_i} , for given $n_i \leq n$, does there exist a presentation

$$e(x) = \sum_{i=1}^m a_i(x) \cdot f_i(x)$$

with analytic functions $a_i(x)$ depending on the same sets of variables as $\widehat{a}_i(x)$? This is the linear version of the nested approximation problem mentioned earlier.

Gabrielov gave in [Ga1] a counterexample to this assertion: Consider the series $f = 1, g = x \cdot (e^z - 1)$, and $h = yz - x$ in three variables x, y, z . He then showed that the convergent series

$$e(x, z) = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \frac{i!}{(i+j)!} \cdot x^i z^{j+1}$$

admits a presentation

$$e = \widehat{a} \cdot f + \widehat{b} \cdot g + \widehat{c} \cdot h,$$

with formal series $\widehat{a}(x, y), \widehat{b}(x, y), \widehat{c}(x, y, z)$ but that there are no convergent series $a(x, y), b(x, y), c(x, y, z)$ representing e in this way. Let us explain why this is so.

It is easy to see that the series $\widehat{a}, \widehat{b}, \widehat{c}$ are unique. Set

$$\widehat{I} = \mathbb{K}[[x, y]] \cdot f + \mathbb{K}[[x, y]] \cdot g + \mathbb{K}[[x, y, z]] \cdot h.$$

Subspaces of $\mathbb{K}[[x_1, \dots, x_n]]$ of the form $\sum_{i=1}^m \mathbb{K}[[x_1, \dots, x_{s_i}]] \cdot f_i$, with $1 \leq s_i \leq n$, extend the notion of ideal in power series rings and are called *echelons* [ACHK].

Now order the monomials $x^i y^j z^k$ lexicographically by their exponents so that $z < y < x$. The *initial monomial* of a nonzero series $s(x, y, z)$ is defined as the smallest monomial $\text{in}(s) = x^i y^j z^k$ of its expansion. We have $\text{in}(f) = 1$, $\text{in}(g) = xz$, and $\text{in}(h) = yz$. The \mathbb{K} -subspace $\text{in}(I)$ of $\mathbb{K}[[x, y, z]]$ spanned by all initial monomials of I turns out to be the sum

$$\text{in}(\widehat{I}) = \mathbb{K}[[x, y]] \cdot 1 + \sum_{k=1}^{\infty} \mathbb{K}[[x]] \cdot x^k z^k + \mathbb{K}[[x, y, z]] \cdot yz.$$

Adapting Buchberger’s algorithm [Bu] for Gröbner bases to the present situation, one constructs series g_2, g_3, \dots in \widehat{I} with initial monomials $x^k z^k$, for $k \geq 2$. The computation gives

$$\begin{aligned} g_2 &= \frac{1}{12} \cdot [x^2 z^2 + \frac{1}{2} \cdot x^2 z^3 + \frac{3}{20} \cdot x^2 z^4 + \frac{1}{30} \cdot x^2 z^5 + \frac{1}{168} \cdot x^2 z^6 + \frac{1}{1120} \cdot x^2 z^7 + \dots], \\ g_3 &= \frac{1}{720} \cdot [x^3 z^3 + \frac{1}{2} \cdot x^3 z^4 + \frac{1}{7} \cdot x^3 z^5 + \frac{5}{168} \cdot x^3 z^6 + \frac{5}{1008} \cdot x^3 z^7 + \dots], \\ g_4 &= \frac{1}{100800} \cdot [x^4 z^4 + \frac{1}{2} \cdot x^4 z^5 + \frac{5}{36} \cdot x^4 z^6 + \frac{1}{36} \cdot x^4 z^7 + \dots], \\ g_5 &= \frac{1}{25401600} \cdot [x^5 z^5 + \frac{1}{2} \cdot x^5 z^6 + \frac{3}{22} \cdot x^5 z^7 + \frac{7}{2461} \cdot x^5 z^8 + \dots], \\ g_6 &= \frac{1}{10059033600} \cdot [x^6 z^6 + \frac{1}{2} \cdot x^6 z^7 + \frac{7}{52} \cdot x^6 z^8 + \frac{1}{39} \cdot x^6 z^9 + \dots]. \end{aligned}$$

The general formula for g_k is

$$g_k = x^k \cdot \sum_{i=k}^{\infty} q_{i,k} \cdot z^i,$$

with coefficients $q_{i,k}$ given by

$$q_{i,k} = \frac{(i-1)!}{4^{k-1} \cdot (i-k)! \cdot (i+k-1)! \cdot (\frac{1}{2})^{\overline{k-1}}},$$

where $(\frac{1}{2})^{\overline{k-1}}$ denotes $\frac{1}{2}(\frac{1}{2} + 1) \cdots (\frac{1}{2} + k - 2)$. The key observation here is that the coefficients

$$q_{k,k} = \frac{(k-1)!}{4^{k-1} \cdot (2k-1)! \cdot (\frac{1}{2})^{\overline{k-1}}}$$

of the initial monomials $x^k z^k$ of g_k *tend very fast* to 0. Rewrite now the series g_k as linear combinations of the original generators f, g, h of \widehat{I} ,

$$g_k = a_k \cdot f + b_k \cdot g + c_k \cdot h,$$

with uniquely defined series $a_k, b_k \in \mathbb{K}[[x, y]]$ and $c_k \in \mathbb{K}[[x, y, z]]$. They are given by the recursions

$$\begin{aligned} a_k &= -y \cdot a_{k-1} + \frac{1}{4(2k-3)(2k-5)} \cdot x^2 \cdot a_{k-2}, \\ b_k &= -y \cdot b_{k-1} + \frac{1}{4(2k-3)(2k-5)} \cdot x^2 \cdot b_{k-2}, \\ c_k &= -y \cdot c_{k-1} + \frac{1}{4(2k-3)(2k-5)} \cdot x^2 \cdot c_{k-2} - z^{-1} \cdot (a_{k-1} \cdot f + b_{k-1} \cdot g + c_{k-1} \cdot h), \end{aligned}$$

with $a_1 = 0$, $a_2 = x^2$, $b_1 = 1$, $b_2 = -y + \frac{1}{2}x$, $c_1 = 0$, $c_2 = -z^{-1} \cdot x \cdot (1 - e^z)$. The preceding formulas imply that a_k, b_k are homogeneous polynomials in x and y

of degree k , respectively $k - 1$, while c_k is a polynomial in x, y, z, e^z, z^{-1} without poles. Note that in the expansions of a_k, b_k , and c_k the monomials x^2y^{k-2}, y^{k-1} , and xy^{k-2} , respectively, appear with coefficients ± 1 . The successive quotients

$$\frac{q_{k+1,k+1}}{q_{k,k}} = \frac{1}{4 \cdot (2k + 1) \cdot (2k - 1)}$$

of the coefficients of g_k tend quadratically towards 0, so the $q_{k,k}$ themselves tend very fast to 0. As

$$\frac{q_{i,k}}{q_{k,k}} = \frac{k! \cdot (i - 1)!}{(i - k)! \cdot (i + k - 1)!} \leq 1$$

for $i \geq k$, all coefficients $q_{i,k}$ of the series g_k become very small as k increases. This then implies that infinite linear combinations of the series g_k with rapidly increasing coefficients may still produce convergent series. A typical example would be the convergent series

$$e(x, z) := \sum_{k=1}^{\infty} \frac{1}{q_{k,k}} \cdot g_k(x, z).$$

By construction, e belongs to $\widehat{I} \cap \mathbb{K}\{x, y, z\}$. We show that it does not belong to I . By uniqueness of the presentation, it suffices to write e as a linear combination $e = \widehat{a} \cdot f + \widehat{b} \cdot g + \widehat{c} \cdot h$ with divergent $\widehat{a}, \widehat{b}, \widehat{c}$. Set $r_k = \frac{1}{q_{k,k}}$ so that $e = \sum_{k=1}^{\infty} r_k \cdot g_k$ and $\widehat{a} = \sum r_k \cdot a_k, \widehat{b} = \sum r_k \cdot b_k, \widehat{c} = \sum r_k \cdot c_k$ with a_k, b_k and c_k as defined above. As we noted earlier, the monomials x^2y^{k-2}, y^{k-1} , and xy^{k-2} appear with coefficients ± 1 in the expansions of a_k, b_k , and c_k , respectively. As the successive quotients r_{k+1}/r_k tend quadratically to infinity, it follows that the series $\widehat{a}, \widehat{b}, \widehat{c}$ diverge. \square

8. TWO APPLICATIONS OF APPROXIMATION

To illustrate the usefulness of the approximation theorems, we give two applications of them to singularity theory.

Application of the analytic and strong approximation theorem. Let \mathbb{K} be a valued field. The group $\text{Aut}(\mathbb{K}\{x\})$ of local \mathbb{K} -algebra automorphisms Φ of $\mathbb{K}\{x\} = \mathbb{K}\{x_1, \dots, x_n\}$ is formed by vectors $\varphi = (\varphi_1, \dots, \varphi_n)$ of convergent power series $\Phi(x_i) = \varphi_i(x) \in (x) \cdot \mathbb{K}\{x\}$ so that the Jacobian matrix $\partial\varphi \in \mathbb{K}\{x\}^{n \times n}$ is invertible, say $\det(\partial\varphi(0)) \neq 0$. The group acts naturally on series $f \in \mathbb{K}\{x\}$ via $\alpha \cdot f = \alpha(f) = f \circ \varphi$. Orbits of a series f consist of all power series g which differ from f by an analytic coordinate change.

Alternatively, we may let the larger group $\mathcal{K} := \mathbb{K}\{x\}^* \rtimes \text{Aut}(\mathbb{K}\{x\})$ act on $\mathbb{K}\{x\}$ via $(u, \Phi) \cdot f = u \cdot \Phi(f) = u \cdot (f \circ \varphi)$. If $f, g \in \mathbb{K}\{x\}$ define hypersurface germs X and Y in \mathbb{K}^n at the origin, then X and Y are isomorphic as germs of analytic spaces if and only if f and g lie in the same \mathcal{K} -orbit.

Now let f_t be an analytic family of convergent power series in $\mathbb{K}\{x\}$, say, a power series $f_t(x) = f(t, x) \in \mathbb{K}\{t, x\}$ in t and x_1, \dots, x_n . One may think of f_t as a family of germs of analytic functions on \mathbb{K}^n at the origin, defined for t varying in a sufficiently small euclidean neighborhood of 0 in \mathbb{K} . Let $(X_t, 0) \subset (\mathbb{K}^n, 0)$ denote the germs of analytic spaces defined by f_t . An important concept in deformation and singularity theory is analytic triviality: Assume that for all t close to 0, the germ $(X_t, 0)$ is isomorphic to the special germ $(X_0, 0)$. This condition is called *pointwise triviality* of the family $(X_t, 0)$. Does this imply that $(X_t, 0)$ is already an *analytically trivial* family in the sense that there exists a family φ_t of analytic

automorphisms of the germ $(\mathbb{K}^n, 0)$ depending analytically on t and sending $(X_t, 0)$ onto $(X_0, 0)$? In terms of equations, this can be rephrased as follows. Assume that f_t belongs to the \mathcal{K} -orbit $\mathcal{K} \cdot f_0$ of f_0 for all t close to 0. Does there exist an analytic family (u_t, φ_t) in \mathcal{K} so that $(u_t, \varphi_t) \cdot f_t = f_0$? By *analytic* we understand here as before that $u_t(x) = u(t, x)$ and $\varphi_t(x) = \varphi(t, x)$ for convergent power series $u \in \mathbb{K}\{t, x\}$ and $\varphi \in \mathbb{K}\{t, x\}^n$.

The problem here is that the condition $f_t \in \mathcal{K} \cdot f_0$ only ensures that, for each t , *some* (u_t, φ_t) exists. But as these are not unique, there is no hope that an arbitrary choice will ensure analytic dependence on the parameter t .

Theorem (Ephraim [Eph, Thm. 0.2], Hauser and Müller [HM2], [HM3]). *Pointwise triviality implies analytic triviality.*

Proof. For $e \in \mathbb{N}$, denote by \mathcal{K}_e the group of e -jets of elements (u, φ) of \mathcal{K} . This means that we truncate u and the components of φ at degree e and consider the action of \mathcal{K}_e on the e -jets $\mathbb{K}[x]_{\leq e}$ of series f in $\mathbb{K}\{x\}$ by taking everything modulo $(x)^{e+1}$. It is easy to see that \mathcal{K}_e is a (finite-dimensional) Lie group over \mathbb{K} which acts on the finite-dimensional vector space $V_e = \mathbb{K}[x]_{\leq e}$. We therefore know from classical differential geometry that the orbits of \mathcal{K}_e are immersed analytic submanifolds of V_e . Moreover, every germ of analytic curve β in V_e lying entirely in one orbit $\mathcal{K}_e \cdot v$ of an element $v \in V_e$ lifts to an analytic curve γ in \mathcal{K}_e inducing β , i.e., so that $\gamma \cdot v = \beta$.

Let us transcribe this statement into equations for our family f_t . We may view f_t as a curve in $V = \mathbb{K}\{x\}$, and then take for β the composition of f_t with the canonical projection $\mathbb{K}\{x\} \rightarrow \mathbb{K}\{x\}/(x)^{e+1}$. As the image of β lies by assumption in one orbit $\mathcal{K}_e \cdot v = \mathcal{K}_e \cdot f(0, x) \subset V_e$, it follows that for every integer e there exists the germ of an analytic curve $\gamma : t \rightarrow (u(t, x), \varphi(t, x)) \in \mathcal{K}_e$ (depending on e) so that

$$u(t, x) \cdot f(t, \varphi(t, x)) \equiv f(0, x) \text{ modulo } (x)^{e+1}.$$

This is an equation which is linear in u and analytic in φ , and which admits a solution for every e . By the strong approximation theorem, there exists a formal exact solution $(\hat{u}(t, x), \hat{\varphi}(t, x)) \in \hat{\mathcal{K}}$,

$$\hat{u}(t, x) \cdot f(t, \hat{\varphi}(t, x)) = f(0, x).$$

Observe that we may require here that \hat{u} and $\hat{\varphi}$ coincide with u and φ up to degree 1, respectively 2, so that \hat{u} is indeed a unit, and $\hat{\varphi}$ is an automorphism. By the analytic approximation theorem, there then even exists an analytic exact solution $(u(t, x), \varphi(t, x)) \in \mathcal{K}$ of the above equation. But the equality

$$u(t, x) \cdot f(t, \varphi(t, x)) = f(0, x)$$

just signifies that the family f_t is analytically trivial. This proves the theorem. \square

Application of the nested algebraic approximation theorem. We consider reduced analytic space germs $(X, 0)$ over a valued field \mathbb{K} . We call such a germ *decomposable* if it is isomorphic to a cartesian product $(X, 0) \cong (Y, 0) \times (Z, 0)$ of two positive-dimensional analytic germs $(Y, 0)$ and $(Z, 0)$. It is clear that every germ admits a factorization into indecomposables, the interesting question is uniqueness of the factors.

Theorem (Hauser and Müller [HM4, Thm. 3]). *Assume that $(X, 0)$ is an analytic space germ in $(\mathbb{K}^n, 0)$ defined by algebraic power series. Then its indecomposable factors are unique up to isomorphism.*

The truly hard part in the proof is the formal case, i.e., the uniqueness of the factorization for spaces defined by formal power series. It does not use any approximation theorem, and we do not give it here. We only show how to deduce via the nested approximation theorem the stated theorem from the formal case.⁴

Proposition. *If an analytic space germ $(X, 0)$ in $(\mathbb{K}^n, 0)$ defined by algebraic power series decomposes formally into a cartesian product $(\widehat{X}, 0) \cong (\widehat{Z}_1, 0) \times (\widehat{Z}_2, 0)$ with $(\widehat{Z}_j, 0)$ formal spaces, then $(X, 0) \cong (Y_1, 0) \times (Y_2, 0)$ itself decomposes with algebraic space germs $(Y_j, 0)$.*

Proof. Let $f_1, \dots, f_p \in \mathbb{K}\langle x \rangle$ define $(X, 0)$ in $(\mathbb{K}^n, 0)$. By the implicit function theorem for algebraic map germs [LT], we may assume that $(X, 0)$ is minimally embedded, i.e., that $f_i \in (x)^2$ for all i . Let J denote the ideal of $\mathbb{K}\langle x \rangle$ generated by the f_i , and denote by $\widehat{J} \subset \mathbb{K}[[x]]$ its completion. By assumption there exists a nontrivial partition of the variables into $x = (x_1, x_2)$ and a formal automorphism $\widehat{\varphi}$ of $\mathbb{K}[[x]]$ so that $\widehat{\Phi}(J)(x) = \widehat{I}_1(x_1) + \widehat{I}_2(x_2)$ for ideals $\widehat{I}_j \subset \mathbb{K}[[x_j]]$. We may assume without loss of generality that $\widehat{\Phi}$ is tangent to the identity, i.e., that $\widehat{\Phi}$ induces the identity on $\mathbb{K}[[x]]/(x)^2$. Write $\widehat{\varphi}$ for the vector in $\mathbb{K}\langle x \rangle^n$ of components $\widehat{\Phi}(x_i)$. The equality is equivalent to the existence of a $(2p \times p)$ -matrix $\widehat{A} = (\widehat{A}_1, \widehat{A}_2)$ of rank p with entries in $\mathbb{K}[[x]]$ so that

$$\begin{aligned} \widehat{A}_1 \cdot f(\widehat{\varphi}(x)) &= f(\widehat{\varphi}(x_1, 0)), \\ \widehat{A}_2 \cdot f(\widehat{\varphi}(x)) &= f(\widehat{\varphi}(0, x_2)). \end{aligned}$$

The system looks like searching for solutions depending on *disjoint* sets of variables, for which approximation is known not to hold [Be]. The loophole is to treat the two equations separately. For each, the nested approximation theorem applies: for the first with respect to $\{x_1\} \subset \{x_1, x_2\}$; for the second with respect to $\{x_2\} \subset \{x_1, x_2\}$. As f is assumed to be algebraic, there exist automorphisms φ_1, φ_2 of $\mathbb{K}\langle x \rangle$, $(p \times p)$ -matrices A_j with entries in $\mathbb{K}\langle x \rangle$, and algebraic power series vectors $\chi_j \in \mathbb{K}\langle x_j \rangle^n$, so that, for $j = 1, 2$,

$$A_j \cdot f(\varphi_j(x)) = f(\chi_j(x_j)).$$

Let ψ_j be the inverses of φ_j , and decompose them into $\psi_j = (\psi_{j1}, \psi_{j2})$ according to $x = (x_1, x_2)$. Then

$$(A_j \circ \psi_j) \cdot f = f(\chi_j \circ \psi_{jj}).$$

As ψ_1 and ψ_2 are tangent to the identity, also (ψ_{11}, ψ_{22}) is an automorphism of $\mathbb{K}\langle x \rangle$. Let ρ be its inverse. Then

$$(A_j \circ \psi_j \circ \rho) \cdot (f \circ \rho) = f \circ \chi_j.$$

But as $f \circ \chi_j$ are algebraic power series vectors in $\mathbb{K}\langle x_j \rangle^p$, we have shown that $(X, 0)$ decomposes into a cartesian product of algebraic space germs. \square

⁴The formal case also implies, via the (nonnested) analytic approximation theorem, that analytic space germs have the cancellation property: if $(X, 0) \times (Z, 0) \cong (Y, 0) \times (Z, 0)$ holds, then already $(X, 0) \cong (Y, 0)$ [HM4, Thm. 1]. Cancellation is a notoriously difficult problem, with much activity in the case of algebraic varieties and biregular or birational morphisms.

9. THE GEOMETRY BEHIND ARTIN'S PROOF

We conclude the article with the transcription of Artin's method of proof to a more differential-geometric setting, as it is developed in detail in [HW], [Wob]. We restrict to the case of one x -variable, which we call t , and one equation f . The starting point is to look at the set of *all* formal power series solutions $\widehat{y}(t)$ of $f(t, y) = 0$, and not just at one particular solution. This set will be considered as an *infinite-dimensional variety* $\widehat{\mathcal{Y}}(f)$. Understanding the geometry of $\widehat{\mathcal{Y}}(f)$ will then allow us to see how formal and convergent, respectively algebraic, solutions are distributed over this variety. This, in turn, will show that for convergent/algebraic f the formal solutions can be approximated in the Krull topology inside $\widehat{\mathcal{Y}}(f)$ by convergent/algebraic ones.

More explicitly, $\widehat{\mathcal{Y}}(f)$ will be stratified into a countable union of locally closed subsets so that each stratum $\widehat{\mathcal{S}}$ is isomorphic (in an appropriate sense) to a cartesian product $Z \times \widehat{M}$ of a finite-dimensional variety Z and a $\mathbb{K}[[t]]$ -module \widehat{M} , both depending on $\widehat{\mathcal{S}}$. Points in Z correspond to polynomial vectors of a prescribed degree (a certain truncation of the formal solutions $\widehat{y}(t)$), and \widehat{M} is shown to be generated by convergent, respectively algebraic, power series whenever f has this quality. Moreover, the second component of the isomorphism towards \widehat{M} sends convergent, respectively, algebraic, power series to power series of the same quality (the first component produces polynomials). The isomorphism thus transfers the comparison of formal, convergent, and algebraic solutions in $\widehat{\mathcal{Y}}(f)$ (more precisely, in $\widehat{\mathcal{S}}$) to a membership problem in \widehat{M} , since the quality is not affected by the component in Z . This is now a linear problem, for which the methods of commutative algebra apply. The analytic and algebraic approximation theorem in one variable follow immediately.

So let $f(t, y) \in \mathbb{K}[[t, y_1, \dots, y_m]]$ be a power series, either formal, convergent, or algebraic. Write the components of vectors $y(t) = (y_1(t), \dots, y_m(t)) \in \mathbb{K}[[t]]^m$ as power series

$$y_j(t) = \sum_{i=0}^{\infty} \alpha_{ij} t^i$$

with coefficients $\alpha_{ij} \in \mathbb{K}$. We have already mentioned in the introduction that, by Taylor expansion and comparison of the coefficients of t^ℓ , the equation $f(t, y(t)) = 0$ induces an equivalent infinite system of polynomial equations

$$(**) \quad F_\ell(x_{ij}) = 0, \quad \ell = 0, 1, \dots$$

in countably many variables x_{ij} , for $j = 1, \dots, m$ and $i \in \mathbb{N}$. Write $\mathbb{K}[x^{\mathbb{N}^m}]$ for the polynomial ring $\mathbb{K}[x_{ij}, j = 1, \dots, m, i \geq 0]$. Then $y(t)$ is a solution of $f(t, y) = 0$ if and only if the coefficients α_{ij} satisfy the system (**). This signifies that the space of formal solutions of $f(t, y) = 0$ is the solution variety $\widehat{\mathcal{Y}}(f)$ of a countable number of polynomial equations inside the infinite-dimensional affine space $\mathbb{A}_{\mathbb{K}}^{\mathbb{N}^m} = (t) \cdot \mathbb{K}[[t]]^m$. The convergent power series form a subspace $\mathbb{B}_{\mathbb{K}}^{\mathbb{N}^m} = (t) \cdot \mathbb{K}\{t\}^m$ inside $\mathbb{A}_{\mathbb{K}}^{\mathbb{N}^m}$, and the analytic version of Artin's theorem aims at comparing $\widehat{\mathcal{Y}}(f)$ with the intersection $\mathcal{Y}(f) = \widehat{\mathcal{Y}}(f) \cap \mathbb{B}_{\mathbb{K}}^{\mathbb{N}^m}$ at a very rudimentary level: if $\widehat{\mathcal{Y}}(f)$ is not void, then also $\mathcal{Y}(f)$ is not void (and, in this case, the latter is dense in the former with respect to the Krull topology).

We go one step further and ask for the actual geometric shape of and for the relation between these two varieties.

So let $y(t)$ be a given formal solution of $f(t, y) = 0$ (to ease the reading, from now on we write $y(t)$ instead of $\widehat{y}(t)$). If all partial derivatives $\partial_{y_i} f$ vanish at $y(t)$, we take them as part of the defining system of equations and start over again with the extended system of equations. Solving several equations requires a more involved argument without any further insight. We will therefore stick to the hypersurface case and assume that some $\partial_{y_i} f(t, y(t))$ is nonzero. By a permutation of the y -variables, we may assume that $\partial_{y_1} f(t, y(t)) \neq 0$. We then have the following.

Theorem (Hauser and Woblistin [HW]). *Associate to a given $f \in \mathbb{K}[[t, y_1, \dots, y_m]]$ its “zero-set”*

$$\mathcal{Y}(f) = \{y(t) \in (t) \cdot \mathbb{K}[[t]]^m, f(t, y(t)) = 0\}$$

inside the space of power series vectors vanishing at 0. Fix $d \in \mathbb{N}$, and set

$$\mathcal{S}_d = \{z(t) \in (t) \cdot \mathbb{K}[[t]]^m, \text{ord}(\partial_{y_1} f(t, z(t))) = d\}.$$

There exists a quasi-affine subvariety $Z \subset \mathbb{A}_{\mathbb{K}}^N$, for some N , and an integer r so that

$$\mathcal{Y}(f) \cap \mathcal{S}_d \cong Z \times \mathbb{K}[[t]]^r.$$

The isomorphism is polynomial in the coefficients of the involved power series and a homeomorphism with respect to the Krull topology. The same statement holds for convergent, respectively algebraic, power series.

Special cases or variants of this result appear at various places in the literature. Denef and Loeser [DL0] treat the case of truncations of arc spaces. Grinberg and Kazhdan [GK] and Drinfeld [Dri] prove a respective factorization for the formal neighborhood of arc spaces. This has been globalized to a certain extent by Bouthier and Kazhdan [BK]; see also [KN, conjecture 73].

Before we start with the proof, let us make some general observations. The stratum \mathcal{S}_d is a locally closed subset of an infinite-dimensional affine space $\mathbb{A}_{\mathbb{K}}^{\mathbb{N}^m}$ in the sense that it is defined by (infinitely many) polynomial equations (=) and inequalities (\neq) in (infinitely many) variables (which correspond to the coefficients of the power series). The series $f \in \mathbb{K}[[t, y]]$ induces a map

$$\begin{aligned} f_{\infty} : (t) \cdot \mathbb{K}[[t]]^m &\rightarrow \mathbb{K}[[t]], \\ z(t) &\rightarrow f(t, z(t)), \end{aligned}$$

given by substitution. The variety $\mathcal{Y}(f)$ is of course just the inverse image $f_{\infty}^{-1}(0)$. We will prove that the restriction of f_{∞} to \mathcal{S}_d can be transformed into a family of affine $\mathbb{K}[[t]]$ -linear maps parametrized by a finite-dimensional algebraic variety. The transformation is carried out by composing f_{∞} with a suitable isomorphism of the source space $(t) \cdot \mathbb{K}[[t]]^m$. This linearization technique—whose present version’s origin lies in Artin’s proof—has been generalized in [BH] to an infinite-dimensional constant rank theorem for maps between power series spaces. It has been exploited successfully in [HW] and [Wob].

The first step consists in dividing the elements $z(t)$ of \mathcal{S}_d by $\partial_{y_1} f(t, z(t))$ according to Weierstrass division. We write henceforth $g(t, y)$ for $\partial_{y_1} f(t, y)$. The division gives

$$z(t) = a(t) \cdot g(t, z(t)) + v(t),$$

where $a(t)$ is a power series that we assume to vanish at 0 (for technical reasons) and where $v(t)$ is a polynomial in t of degree $\leq d$. A simple argument (see the lemma below) shows that $g(t, z(t))$ and $g(t, v(t))$ generate the same ideal in $\mathbb{K}[[t]]$. Hence $v(t)$ again belongs to \mathcal{S}_d . Moreover, the coefficients of $v(t)$ are polynomials in the coefficients of $z(t)$.

In what follows, we just write v and z for $z(t)$ and $v(t)$ for short.

Lemma ([HW, Prop. 5.2]). *Let $\mathcal{V}_d = (t) \cdot \mathbb{K}[t]_{\leq d}^m$ denote the space of polynomial vectors of degree $\leq d$ vanishing at 0. The map*

$$\begin{aligned} \varphi_d : (t) \cdot \mathbb{K}[[t]]^m \times (\mathcal{V}_d \cap \mathcal{S}_d) &\rightarrow \mathcal{S}_d, \\ (a, v) &\rightarrow z = v + a \cdot g(t, v) \end{aligned}$$

is bijective and is a homeomorphism for the Krull topology. The same statement holds in the convergent and algebraic case.

Notice that \mathcal{V}_d is a finite-dimensional \mathbb{K} -vector space and consequently $\mathcal{V}_d \cap \mathcal{S}_d$ is a quasi-affine variety.

Proof. We first show that φ_d is well-defined, i.e., has its image in \mathcal{S}_d . For this we have to show that $g(t, z)$ has again order d . But, by binomial expansion, $g(t, z) = g(t, v + a \cdot g(t, v)) = u(t) \cdot g(t, v)$ holds for some unit $u \in \mathbb{K}[[t]]^*$. In particular, $g(t, z)$ and $g(t, v)$ generate the same ideals of $\mathbb{K}[[t]]$. Hence $g(t, z)$ has order d , i.e., $z \in \mathcal{S}_d$.

It is then clear that φ_d is continuous. The inverse to φ_d is given by Weierstrass division in each component of $z \in \mathcal{S}_d$. Namely, as $\text{ord } g(t, z) = d$, each z_i can be written in a unique way as $z_i = b_i \cdot g(t, z) + v_i$ with $b_i \in (t) \cdot \mathbb{K}[[t]]$ and $v_i \in (t) \cdot \mathbb{K}[t]$ a polynomial of degree $\leq d$. Set $v = (v_1, \dots, v_m) \in (t) \cdot \mathbb{K}[t]_{\leq d}^m$. By the same argument as before, we have $g(t, z) = u(t) \cdot g(t, v)$ for some unit $u \in \mathbb{K}[[t]]^*$. Setting $a = (a_1, \dots, a_m)$ with $a_i = u(t) \cdot b_i$, we get $z = v + a \cdot g(t, v)$ as required. As the Weierstrass division is continuous; φ_d^{-1} is also. This gives the lemma. \square

Proof of the Theorem. Composing f_∞ with φ_d , we get a map

$$\begin{aligned} h_\infty := f_\infty \circ \varphi_d : (t) \cdot \mathbb{K}[[t]]^m \times (\mathcal{V}_d \cap \mathcal{S}_d) &\rightarrow \mathbb{K}[[t]], \\ (a, v) &\rightarrow f(t, v + a \cdot g(t, v)), \end{aligned}$$

induced by $h(t, y) = f(t, y + a \cdot g(t, y)) \in \mathbb{K}[[t, y]]$. This h is the power series composition of f with $v + a \cdot g(t, v)$. By construction, φ_d^{-1} maps $\mathcal{Y}(f) \cap \mathcal{S}_d$ homeomorphically onto

$$\mathcal{Y}(h) \cap [(t) \cdot \mathbb{K}[[t]]^m \times (\mathcal{V}_d \cap \mathcal{S}_d)] = h_\infty^{-1}(0).$$

The clue now is to show that this zero-set is homeomorphic to an affine $\mathbb{K}[[t]]$ -module (i.e., the translate of a $\mathbb{K}[[t]]$ -module). But this is easy. Indeed, we may again apply Taylor expansion to $f(t, v + a \cdot g(t, v))$. It gives

$$h(t, a, v) = f(t, v + a \cdot g(t, v)) = f(t, v) + \partial_y f(t, v) \cdot a \cdot g(t, v) + \dots,$$

where the dots collect the terms which are quadratic or of higher degree in the components of $a \cdot g(t, v)$. Recall that $g = \partial_{y_1} f$. So we may rewrite the last formula as

$$h(t, a, v) = f(t, v) + (a_1 + g(t, a)) \cdot g(t, v)^2 + \sum_{i=2}^m \partial_{y_i} f(t, v) \cdot a_i \cdot g(t, v),$$

where $q(t, a) \cdot g(t, v)^2$ collects all higher-order terms of the expansion, and $q(t, a)$ is at least quadratic in the components of a . By the inverse function theorem, the map

$$\begin{aligned} \chi &: (t) \cdot \mathbb{K}[[t]]^m \rightarrow (t) \cdot \mathbb{K}[[t]]^m, \\ a &\rightarrow (a_1 + q(t, a), a_2, \dots, a_m) \end{aligned}$$

is a (global) isomorphism (it is here that we use that $a(t)$ vanishes at 0). By definition, $h(t, a, v) = k(t, \chi(a), v)$, with k the power series associated to the substitution map

$$\begin{aligned} k_\infty &: (t) \cdot \mathbb{K}[[t]]^m \times (\mathcal{V}_d \cap \mathcal{S}_d) \rightarrow \mathbb{K}[[t]], \\ (a, v) &\rightarrow f(t, v) + a_1 \cdot g(t, v)^2 + \sum_{i=2}^m \partial_{y_i} f(t, v) \cdot a_i \cdot g(t, v). \end{aligned}$$

If we write $e(t, v)$ for the vector $g(t, v) \cdot \partial_y f(t, v) \in \mathbb{K}[[t]]^m$, this reads more clearly as

$$k_\infty(a, v) = f(t, v) + a \cdot e(t, v),$$

where $a \cdot e(t, v)$ denotes the scalar product. The map k_∞ is hence linear in a , and the coefficients of the image series are polynomials in the coefficients of v .

As $(a, v) \rightarrow (\chi(a), v)$ is a homeomorphism, it follows that $\mathcal{Y}(f) \cap \mathcal{S}_d$ is homeomorphic to the kernel \mathcal{U}_d of k_∞ . Consider now the projection $\mathcal{U}_d \rightarrow \mathcal{V}_d \cap \mathcal{S}_d$ on the second factor. This defines an ‘‘affine $\mathbb{K}[[t]]$ -submodule bundle’’ over the finite-dimensional variety $\mathcal{V}_d \cap \mathcal{S}_d$. By this we understand that the fibers over $v \in \mathcal{V}_d \cap \mathcal{S}_d$ are affine $\mathbb{K}[[t]]$ -modules $\mathcal{U}_{d,v}$ given as the solution spaces of an inhomogeneous $\mathbb{K}[[t]]$ -linear equation, namely

$$\mathcal{U}_{d,v} = \{a \in (t) \cdot \mathbb{K}[[t]]^m, a \cdot e(t, v) = -f(t, v)\}.$$

The fibers are nonempty if and only if $f(t, v)$ belongs to the ideal $J(t, v)$ of $\mathbb{K}[[t]]$ generated by the series $\partial_{y_i} f(t, v) \cdot \partial_{y_i} f(t, v)$, for $i = 1, \dots, m$. Now define $Z \subset \mathcal{V}_d \cap \mathcal{S}_d$ as the vectors v satisfying this membership condition. Via the Weierstrass theorem (this time with parameters), one can show that Z is a (closed) algebraic subvariety of $\mathcal{V}_d \cap \mathcal{S}_d$, given by setting the remainder of the division of $f(t, v)$ by the ideal $J(t, v)$ equal to zero.

This bundle structure of \mathcal{U}_d is not yet the required cartesian product. It can be obtained by a slight modification of the division used in the preceding lemma.

Lemma’. *Let $\mathcal{V}'_d = (t) \cdot \mathbb{K}[t]_{\leq d} \times (t) \cdot \mathbb{K}[t]_{\leq 2d}^{m-1}$ denote the space of polynomial vectors vanishing at 0 of degree $\leq d$ in the first component, and of degree $\leq 2d$ in the remaining components. The map*

$$\varphi'_d : (t) \cdot \mathbb{K}[[t]]^m \times (\mathcal{V}'_d \cap \mathcal{S}_d) \rightarrow \mathcal{S}_d,$$

given componentwise by

$$\begin{aligned} (a_1, v_1) &\rightarrow z_1 = v_1 + a_1 \cdot g(t, v), \\ (a_i, v_i) &\rightarrow z_i = v_i + a_i \cdot g(t, v)^2, \end{aligned}$$

for $i \geq 2$, is bijective and a homeomorphism for the Krull topology. The same statement holds in the convergent and algebraic case.

Proof of Lemma'. The same argument as for the original lemma shows that φ'_d is well-defined. The inverse to φ'_d is obtained by dividing the first component of a vector $z(t)$ by $g(t, z(t))$, and the other components by $g(t, z(t))^2$. This gives the claim. \square

We finish the proof of the theorem. Using the modified isomorphism φ'_d instead of φ_d yields, by the same arguments as before, the following equation defining the modified \mathcal{U}_d :

$$f(t, v) + (a_1 + \sum_{i=2}^m \partial_{y_i} f(t, v) \cdot a_i) \cdot g(t, v)^2 = 0.$$

We compose with the isomorphism

$$a \rightarrow (a_1 - \sum_{i=2}^m \partial_{y_i} f(t, v) \cdot a_i, a_2, \dots, a_m),$$

and we get a new set \mathcal{U}'_d defined by

$$f(t, v) + a_1 \cdot g(t, v)^2 = 0.$$

This prescribes a_1 uniquely for v varying in the set $Z' \subset \mathcal{V}'_d \cap \mathcal{S}_d$ defined by

$$Z' = \{v \in \mathcal{V}'_d \cap \mathcal{S}_d, f(t, v) \in (g(t, v)^2)\},$$

where $(g(t, v)^2)$ denotes the ideal of $\mathbb{K}[[t]]$ generated by $g(t, v)^2$. Indeed, for these v we have $a_1 = -f(t, v)/g(t, v)^2$. From all this we conclude that

$$\mathcal{Y}(f) \cap \mathcal{S}_d \cong Z' \times (t) \cdot \mathbb{K}[[t]]^{m-1}.$$

This is the product decomposition asserted in the theorem. \square

Remark. From a differential-geometric viewpoint, the preceding proof admits the following interpretation. The first version of the lemma yields a decomposition of power series vectors $y(t)$ into pairs $(v(t), a(t))$ so that the map f_∞ becomes a map of constant rank in a , for each v . Composing it with a suitable isomorphism, its fibers are shown to form a bundle over the space of v 's. The second lemma defines a more refined decomposition which directly produces a trivial bundle, i.e., a cartesian product.

From the factorization theorem we immediately get the following.

Corollary. *In one variable t , analytic and algebraic Artin approximation holds.*

Proof. All isomorphisms appearing in the preceding construction restrict equally to the spaces of convergent or algebraic power series. The second component v is not affected by these restrictions since the components of v are polynomials. The existence of a formal solution $\widehat{y}(t)$ of $f(t, y) = 0$ is equivalent, by the above, to saying that the image $(\widehat{a}(t), \widehat{v}(t))$ of $\widehat{y}(t)$ has second component $\widehat{v}(t)$ so that the fiber of $\widehat{\mathcal{U}}_d$ above $\widehat{v}(t)$ is nonempty. But $\widehat{v}(t)$ is a polynomial, so we may equivalently write $v(t)$ for $\widehat{v}(t)$. As the fiber above $v(t)$ is an ideal generated by convergent, respectively algebraic, power series when $f(t, y)$ is convergent, respectively algebraic, the intersection of the fiber with the space of convergent, respectively algebraic, power series is also nonempty. So we find a convergent $a(t)$ in the fiber above $v(t)$. Going backwards, the convergent solution $y(t)$ to $f(t, y) = 0$ is found. As all maps are continuous in the Krull topology, we also get the density statement. \square

ACKNOWLEDGMENTS

The work on the article at hand has profited very much from the special semester “Artin Approximation and Infinite Dimensional Geometry” within the Chaire Jean Morlet of the author at CIRM, Luminy-Marseille, in the spring of 2015. The local co-organizer was Guillaume Rond. The excellent working conditions and the kind hospitality of the staff at CIRM are highly appreciated. The success of the semester owns very much to the enthusiasm of the participants.

Many people contributed to the drafting of the present text, either by discussions with the author or by their own work on approximation. Among them we would like to mention the following: M. Artin, B. Adamczewski, M. E. Alonso, A. Bostan, M. Bousquet-Mélou, F. J. Castro-Jiménez, C. Chiu, A. Gioia, M. Hickel, S. Izumi, C. Koutschan, B. Lamel, L. Lipshitz, S. Mori, G. Pfister, A. Płoski, D. Popescu, K. Raschel, M. Roczen, G. Rond, K. Slavov, M. Spivakovsky, B. Teissier, and S. Woblistin.

The suggestions of an anonymous and very competent referee helped to improve and clarify various arguments in the presentation.

ABOUT THE AUTHOR

Herwig Hauser is professor of mathematics at the University of Vienna, Austria. His research focuses on singularities of algebraic varieties, especially their resolution, and on problems related to Artin approximation and infinite-dimensional varieties of power series spaces. He is also the author of many visualizations of algebraic surfaces, known through the exposition series “Imaginary” and the movie “Zeroset—I spy with my little eye”.

REFERENCES

- [Abe] N. H. Abel, *Œuvres complètes. Tome I* (French), Éditions Jacques Gabay, Sceaux, 1992. Edited and with a preface by L. Sylow and S. Lie; Reprint of the second (1881) edition. MR1191901
- [ACH] M. E. Alonso, F. J. Castro-Jiménez, H. Hauser, *Encoding algebraic power series*. Found. Comp. Math. 2017. To appear.
- [ACHK] M. E. Alonso, F. J. Castro-Jiménez, H. Hauser, C. Koutschan, *Echelons of power series and Gabrielov’s counterexample to nested linear Artin approximation*. Manuscript 2017, 8 pp.
- [AB] B. Adamczewski and J. P. Bell, *Diagonalization and rationalization of algebraic Laurent series* (English, with English and French summaries), Ann. Sci. Éc. Norm. Supér. (4) **46** (2013), no. 6, 963–1004, DOI 10.24033/asens.2207. MR3134685
- [AM] M. Artin and B. Mazur, *On periodic points*, Ann. of Math. (2) **81** (1965), 82–99, DOI 10.2307/1970384. MR0176482
- [AMD] M. Atiyah, I. MacDonald, *Introduction to commutative algebra*, Addison-Wesley Publishing Co., Reading, MA, 1969. MR0242802 (39 #4129)
- [AMR] M. E. Alonso, T. Mora, and M. Raimondo, *A computational model for algebraic power series*, J. Pure Appl. Algebra **77** (1992), no. 1, 1–38, DOI 10.1016/0022-4049(92)90029-F. MR1148270
- [Ar1] M. Artin, *On the solutions of analytic equations*, Invent. Math. **5** (1968), 277–291, DOI 10.1007/BF01389777. MR0232018
- [Ar2] M. Artin, *Algebraic approximation of structures over complete local rings*, Inst. Hautes Études Sci. Publ. Math. **36** (1969), 23–58. MR0268188
- [Be] J. Becker, *A counterexample to Artin approximation with respect to subrings*, Math. Ann. **230** (1977), no. 2, 195–196, DOI 10.1007/BF01370664. MR0480508

- [BDLv] J. Becker, J. Denef, L. Lipshitz, and L. van den Dries, *Ultraproducts and approximations in local rings. I*, *Invent. Math.* **51** (1979), no. 2, 189–203, DOI 10.1007/BF01390228. MR528023
- [BD] C. Banderier and M. Drmota, *Formulae and asymptotics for coefficients of algebraic functions*, *Combin. Probab. Comput.* **24** (2015), no. 1, 1–53, DOI 10.1017/S0963548314000728. MR3318039
- [BH] C. Bruschek and H. Hauser, *Arcs, cords, and felts—six instances of the linearization principle*, *Amer. J. Math.* **132** (2010), no. 4, 941–986, DOI 10.1353/ajm.0.0134. MR2663645
- [BK] A. Bouthier, D. Kazhdan, *Faisceaux pervers sur les espaces d’arcs I: Le cas d’égales caractéristiques*. arXiv: 1509.02203v2.
- [Bos] A. Bostan, *Algebraicity and transcendence of power series: combinatorial and computational aspects*. Notes from lectures at the University of Linz, 2016.
- [Bou] N. Bourbaki, *Commutative Algebra*, Springer, 1989.
- [Bu] B. Buchberger, *Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems* (German), *Aequationes Math.* **4** (1970), 374–383, DOI 10.1007/BF01844169. MR0268178
- [Chr] G. Christol, *Diagonales de fractions rationnelles* (French), *Séminaire de Théorie des Nombres, Paris 1986–87*, *Progr. Math.*, vol. 75, Birkhäuser Boston, Boston, MA, 1988, pp. 65–90. MR990506
- [Coc] J. Cockle, *On transcendental and algebraic solution*. *Philosophical Magazine XXI* (1861), 379–383.
- [CPR] F. C. Castro-Jiménez, D. Popescu, G. Rond, *Linear nested Artin approximation theorem for algebraic power series*. Preprint 2016.
- [Del] P. Deligne, *Intégration sur un cycle évanescant* (French), *Invent. Math.* **76** (1984), no. 1, 129–143, DOI 10.1007/BF01388496. MR739629
- [dJP] T. de Jong and G. Pfister, *Local analytic geometry: Basic theory and applications*, *Advanced Lectures in Mathematics*, Friedr. Vieweg & Sohn, Braunschweig, 2000. MR1760953
- [Den] J. Denef, *The Diophantine problem for polynomial rings and fields of rational functions*, *Trans. Amer. Math. Soc.* **242**, posted on 1978, 391–399, DOI 10.2307/1997746. MR0491583
- [DL1] J. Denef and L. Lipshitz, *Ultraproducts and approximation in local rings. II*, *Math. Ann.* **253** (1980), no. 1, 1–28, DOI 10.1007/BF01457817. MR594530
- [DL2] J. Denef and L. Lipshitz, *Power series solutions of algebraic differential equations*, *Math. Ann.* **267** (1984), no. 2, 213–238, DOI 10.1007/BF01579200. MR738249
- [DL3] J. Denef and L. Lipshitz, *Algebraic power series and diagonals*, *J. Number Theory* **26** (1987), no. 1, 46–67, DOI 10.1016/0022-314X(87)90095-3. MR883533
- [DLo] J. Denef and F. Loeser, *Germes of arcs on singular algebraic varieties and motivic integration*, *Invent. Math.* **135** (1999), no. 1, 201–232, DOI 10.1007/s002220050284. MR1664700
- [Dri] V. Drinfeld: On the Grinberg-Kazhdan formal arc theorem. arXiv:math/0203263.
- [dV] L. Di Vizio, *Arithmetic theory of q -difference equations: the q -analogue of Grothendieck-Katz’s conjecture on p -curvatures*, *Invent. Math.* **150** (2002), no. 3, 517–578, DOI 10.1007/s00222-002-0241-z. MR1946552
- [dVRSZ] L. Di Vizio, J.-P. Ramis, J. Sauloy, and C. Zhang, *Équations aux q -différences* (French), *Gaz. Math.* **96** (2003), 20–49. MR1988639
- [EGA] A. Grothendieck and J. A. Dieudonné, *Eléments de géométrie algébrique. I* (French), *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 166, Springer-Verlag, Berlin, 1971. MR3075000
- [Eis] G. Eisenstein, *Über eine allgemeine Eigenschaft der Reihen-Entwicklungen aller algebraischen Funktionen*. *Berichte Königl. Preuss. Akad. Wiss. Berlin* (1852), 441–443. Reproduced in: *Mathematische Gesammelte Werke, Band II*, Chelsea Publishing 1975, pp. 765–767.
- [Eph] R. Ephraim, *Isosingular loci and the Cartesian product structure of complex analytic singularities*, *Trans. Amer. Math. Soc.* **241** (1978), 357–371, DOI 10.2307/1998850. MR492307

- [Fur] H. Furstenberg, *Algebraic functions over finite fields*, J. Algebra **7** (1967), 271–277, DOI 10.1016/0021-8693(67)90061-0. MR0215820
- [Ga1] A. M. Gabrièlov, *The formal relations between analytic functions* (Russian), Funkcional. Anal. i Priložen. **5** (1971), no. 4, 64–65. MR0302930
- [Ga2] A. M. Gabrièlov, *Formal relations among analytic functions* (Russian), Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 1056–1090. MR0346184
- [Gal] A. Galligo, *Théorème de division et stabilité en géométrie analytique locale* (French, with English summary), Ann. Inst. Fourier (Grenoble) **29** (1979), no. 2, vii, 107–184. MR539695
- [GK] M. Grinberg and D. Kazhdan, *Versal deformations of formal arcs*, Geom. Funct. Anal. **10** (2000), no. 3, 543–555, DOI 10.1007/PL00001628. MR1779611
- [Gra] H. Grauert, *Über die Deformation isolierter Singularitäten analytischer Mengen* (German), Invent. Math. **15** (1972), 171–198, DOI 10.1007/BF01404124. MR0293127
- [Gre] M. J. Greenberg, *Rational points in Henselian discrete valuation rings*, Inst. Hautes Études Sci. Publ. Math. **31** (1966), 59–64. MR0207700
- [Gu] R. V. Gurjar, *A hyperplane section theorem for Milnor numbers*, Math. Ann. **309** (1997), no. 2, 341–357, DOI 10.1007/s002080050116. MR1474196
- [Hara] T. Harase, *Algebraic elements in formal power series rings*, Israel J. Math. **63** (1988), no. 3, 281–288, DOI 10.1007/BF02778035. MR969943
- [Har1] R. Harley, *On the theory of the transcendental solution of algebraic equations*. Quart. J. Pure Applied Math. **5** (1862), 337–361.
- [HM1] H. Hauser and G. Müller, *A rank theorem for analytic maps between power series spaces*, Inst. Hautes Études Sci. Publ. Math. **80** (1994), 95–115 (1995). MR1320605
- [HM2] H. Hauser and G. Müller, *Analytic curves in power series rings*, Compositio Math. **76** (1990), no. 1-2, 197–201. Algebraic geometry (Berlin, 1988). MR1078862
- [HM3] H. Hauser and G. Müller, *The trivial locus of an analytic map germ* (English, with French summary), Ann. Inst. Fourier (Grenoble) **39** (1989), no. 4, 831–844. MR1036334
- [HM4] H. Hauser and G. Müller, *The cancellation property for direct products of analytic space germs*, Math. Ann. **286** (1990), no. 1-3, 209–223, DOI 10.1007/BF01453573. MR1032931
- [HW] H. Hauser, S. Woblistin, *On the structure of varieties of power series in one variable*. Manuscript 2016, 39 pp.
- [Hei] E. Heine, *Der Eisensteinsche Satz über Reihen-Entwicklung algebraischer Functionen* (German), J. Reine Angew. Math. **45** (1853), 285–302, DOI 10.1515/crll.1853.45.285. MR1578830
- [Her] G. Hermann, *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale* (German), Math. Ann. **95** (1926), no. 1, 736–788, DOI 10.1007/BF01206635. MR1512302
- [Hic] M. Hickel, *Fonction de Artin et germes de courbes tracées sur un germe d’espace analytique* (French, with French summary), Amer. J. Math. **115** (1993), no. 6, 1299–1334, DOI 10.2307/2374967. MR1254735
- [Hir] H. Hironaka, *Idealistic exponents of singularity*, Algebraic geometry (J. J. Sylvester Sympos., Johns Hopkins Univ., Baltimore, Md., 1976), Johns Hopkins Univ. Press, Baltimore, Md., 1977, pp. 52–125. MR0498562
- [HiRo] H. Hironaka and H. Rossi, *On the equivalence of imbeddings of exceptional complex spaces*, Math. Ann. **156** (1964), 313–333, DOI 10.1007/BF01361027. MR0171784
- [Izu] S. Izumi, *Convergence of formal morphisms of completions of complex spaces*, J. Math. Soc. Japan **51** (1999), no. 3, 731–755, DOI 10.2969/jmsj/05130731. MR1691457
- [KN] J. Kollár and A. Némethi, *Holomorphic arcs on singularities*, Invent. Math. **200** (2015), no. 1, 97–147, DOI 10.1007/s00222-014-0530-3. MR3323576
- [Kat1] N. M. Katz, *Algebraic solutions of differential equations (p -curvature and the Hodge filtration)*, Invent. Math. **18** (1972), 1–118, DOI 10.1007/BF01389714. MR0337959
- [Kat2] N. M. Katz, *A conjecture in the arithmetic theory of differential equations* (English, with French summary), Bull. Soc. Math. France **110** (1982), no. 2, 203–239. MR667751
- [KMPPR] H. Kurke, T. Mostowski, G. Pfister, D. Popescu, and M. Roczen, *Die Approximationseigenschaft lokaler Ringe* (German), Lecture Notes in Mathematics, Vol. 634, Springer-Verlag, Berlin-New York, 1978. MR0485851

- [Laf1] J.-P. Lafon, *Séries formelles algébriques* (French), C. R. Acad. Sci. Paris **260** (1965), 3238–3241. MR0174585
- [Laf2] J.-P. Lafon, *Anneaux henséliens et théorème de préparation* (French), C. R. Acad. Sci. Paris Sér. A-B **264** (1967), A1161–A1162. MR0214589
- [Lan] S. Lang, *Some applications of the local uniformization theorem*, Amer. J. Math. **76** (1954), 362–374, DOI 10.2307/2372578. MR0062722
- [Lip] L. Lipshitz, *D-finite power series*, J. Algebra **122** (1989), no. 2, 353–373, DOI 10.1016/0021-8693(89)90222-6. MR999079
- [LT] F. Lazzeri and A. Tognoli, *Alcune proprietà degli spazi algebrici* (Italian), Ann. Scuola Norm. Sup. Pisa (3) **24** (1970), 597–632. MR0292827
- [Mu] D. Mumford, *The red book of varieties and schemes*, Second, expanded edition, Lecture Notes in Mathematics, vol. 1358, Springer-Verlag, Berlin, 1999. MR1748380
- [Ma] H. Matsumura, *Commutative algebra*, 2nd ed., Mathematics Lecture Note Series, vol. 56, Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1980. MR575344
- [Mo] S. Mori, *On a hyperplane section theorem of Gurjar*, Math. Ann. **319** (2001), no. 3, 533–537, DOI 10.1007/PL00004447. MR1819882
- [Na] M. Nagata, *On the theory of Henselian rings*, Nagoya Math. J. **5** (1953), 45–57. MR0051821
- [Og] T. Ogoma, *General Néron desingularization based on the idea of Popescu*, J. Algebra **167** (1994), no. 1, 57–84, DOI 10.1006/jabr.1994.1175. MR1282816
- [Os] W. F. Osgood, *On functions of several complex variables*, Trans. Amer. Math. Soc. **17** (1916), no. 1, 1–8, DOI 10.2307/1988823. MR1501027
- [PfpPo] G. Pfister and D. Popescu, *Die strenge Approximationseigenschaft lokaler Ringe* (German), Invent. Math. **30** (1975), no. 2, 145–174, DOI 10.1007/BF01425506. MR0379490
- [P11] A. Płoski, *Note on a theorem of M. Artin* (English, with Russian summary), Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **22** (1974), 1107–1109. MR0367257
- [P12] A. Płoski, *Les solutions formelles et convergentes des équations analytiques*. Semestre de Singularités à Lille, 1999.
- [P13] A. Płoski, *Formal and convergent solutions of analytic equations*. Manuscript 2015.
- [Po1] D. Popescu, *General Néron desingularization*, Nagoya Math. J. **100** (1985), 97–126. MR818160
- [Po2] D. Popescu, *General Néron desingularization and approximation*, Nagoya Math. J. **104** (1986), 85–115. MR868439
- [Po3] D. Popescu, *A strong approximation theorem over discrete valuation rings*, Rev. Roumaine Math. Pures Appl. **20** (1975), no. 6, 659–692. MR0393023
- [Po4] D. Popescu, *Artin approximation*, Handbook of algebra, Vol. 2, Handb. Algebr., vol. 2, Elsevier/North-Holland, Amsterdam, 2000, pp. 321–356, DOI 10.1016/S1570-7954(00)80031-5. MR1759599
- [Pol] G. Pólya, *Sur les séries entières, dont la somme est une fonction algébrique*. L'Enseign. Math. **22** (1921-22), 38-47.
- [Qu] R. Quarez, *The Artin conjecture for \mathbf{Q} -algebras*, Rev. Mat. Univ. Complut. Madrid **10** (1997), no. 2, 229–263. MR1605646
- [Ra1] M. Raynaud, *Travaux récents de M. Artin* (French), Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363, Lecture Notes in Math., vol. 175, Springer, Berlin, 1971, pp. Exp. No. 363, 279–295. MR3077132
- [Ra2] M. Raynaud, *Anneaux locaux henséliens* (French), Lecture Notes in Mathematics, Vol. 169, Springer-Verlag, Berlin-New York, 1970. MR0277519
- [Ron1] G. Rond, *Sur la linéarité de la fonction de Artin* (French, with English and French summaries), Ann. Sci. École Norm. Sup. (4) **38** (2005), no. 6, 979–988, DOI 10.1016/j.ansens.2005.10.001. MR2216836
- [Ron2] G. Rond, *Lemme d'Artin-Rees, théorème d'Izumi et fonction de Artin* (French, with English summary), J. Algebra **299** (2006), no. 1, 245–275, DOI 10.1016/j.jalgebra.2005.06.037. MR2225775
- [Rot] C. Rotthaus, *On the approximation property of excellent rings*, Invent. Math. **88** (1987), no. 1, 39–63, DOI 10.1007/BF01405090. MR877005
- [Rui] J. M. Ruiz, *The basic theory of power series*, Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1993. MR1234937

- [ShW] H. Sharif and C. F. Woodcock, *Algebraic functions over a field of positive characteristic and Hadamard products*, J. London Math. Soc. (2) **37** (1988), no. 3, 395–403, DOI 10.1112/jlms/s2-37.3.395. MR939116
- [Spa] K. Spallek, *Differenzierbare Kurven auf analytischen Mengen* (German), Math. Ann. **177** (1968), 54–66, DOI 10.1007/BF01350730. MR0228004
- [Spi1] M. Spivakovsky, *Non-existence of the Artin function for Henselian pairs*, Math. Ann. **299** (1994), no. 4, 727–729, DOI 10.1007/BF01459808. MR1286894
- [Spi2] M. Spivakovsky, *A new proof of D. Popescu’s theorem on smoothing of ring homomorphisms*, J. Amer. Math. Soc. **12** (1999), no. 2, 381–444, DOI 10.1090/S0894-0347-99-00294-5. MR1647069
- [Sta] R. P. Stanley, *Differentiably finite power series*, European J. Combin. **1** (1980), no. 2, 175–188, DOI 10.1016/S0195-6698(80)80051-5. MR587530
- [Sw] R. G. Swan, *Néron-Popescu desingularization*, Algebra and geometry (Taipei, 1995), Lect. Algebra Geom., vol. 2, Int. Press, Cambridge, MA, 1998, pp. 135–192. MR1697953
- [Tan] J. Tannery, *Propriétés des intégrales des équations différentielles linéaires à coefficients variables*. Thèse de doctorat, Faculté des Sciences Paris 1874. Available at <http://gallica.bnf.fr>.
- [Te] B. Teissier, *Résultats récents sur l’approximation des morphismes en algèbre commutative (d’après André, Artin, Popescu et Spivakovsky)* (French, with French summary), Astérisque **227** (1995), Exp. No. 784, 4, 259–282. Séminaire Bourbaki, Vol. 1993/94. MR1321650
- [Tou] J.-C. Tougeron, *Idéaux de fonctions différentiables*, Springer-Verlag, Berlin-New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 71. MR0440598
- [Wa] J. J. Wavrik, *A theorem on solutions of analytic equations with applications to deformations of complex structures*, Math. Ann. **216** (1975), no. 2, 127–142, DOI 10.1007/BF01432540. MR0387649
- [Wob] S. Woblistin, *On varieties of power series spaces*. PhD thesis, University of Vienna 2016.
- [Za] O. Zariski, *Analytical irreducibility of normal varieties*, Ann. of Math. (2) **49** (1948), 352–361, DOI 10.2307/1969284. MR0024158
- [ZS] O. Zariski and P. Samuel, *Commutative algebra. Vol. II*, Springer-Verlag, New York-Heidelberg, 1975. Reprint of the 1960 edition; Graduate Texts in Mathematics, Vol. 29. MR0389876

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, AUSTRIA
E-mail address: herwig.hauser@univie.ac.at